SCHOOL OF EDUCATION
LIBRARY

Volume VII
Number 7

SEPTEMBER, 1916

$2.50 a Year
30c. Per Copy

THE JOURNAL
OF
Educational Psychology

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

EDITORS:

W. C. BAGLEY,
University of Illinois.

C. E. SEASHORE,
University of Iowa.

J. CARLETON BELL, Managing Editor,
Brooklyn Training School for Teachers.

GUY MONTROSE WHIPPLE,
University of Illinois.

PUBLISHED MONTHLY, EXCEPT JULY AND AUGUST.
WARWICK & YORK, INC.,
LANCASTER, PA. BALTIMORE, MD.

[Entered as second-class matter at the postoffice at Lancaster, Pa., under the Act of March 3, 1879.]

COLLABORATORS

SYDNEY ALRUTZ,
University of Upsala, Sweden.

J. R. ANGELL,
University of Chicago.

J. W. BAIRD,
Clark University.

BIRD T. BALDWIN,
Swarthmore College.

W. VAN DYKE BINGHAM,
Carnegie Institute of Technology.

E. F. BUCHNER,
Johns Hopkins University.

W. H. BURNHAM,
Clark University.

ED. CLAPARÈDE,
University of Geneva, Switzerland.

S. S. COLVIN,
Brown University.

W. F. DEARBORN,
Harvard University.

H. H. GODDARD,
New Jersey Training School, Vineland, N. J.

JOSEPH JASTROW,
University of Wisconsin.

G. E. JOHNSON,
Harvard University.

E. A. KIRKPATRICK,
State Normal School, Fitchburg, Mass.

E. H. LINDLEY,
University of Indiana.

ANNA J. MCKEAG,
Wellesley College.

D. P. MACMILLAN,
Board of Education, Chicago, Ill.

D. R. MAJOR,
Columbus, Ohio.

ADOLF MEYER,
Johns Hopkins University.

J. B. MINER,
Carnegie Institute of Technology.

WILL S. MONROE,
State Normal School, Montclair, N. J.

E. C. MOORE,
Harvard University.

HUGO MÜNSTERBERG,
Harvard University.

M. V. O'SHEA,
University of Wisconsin.

W. B. PILLSBURY,
University of Michigan.

S. H. ROWE,
Principal Wadleigh High School, New York City.

SANTE DE SANCTIS,
University of Rome, Italy.

E. C. SANFORD,
President of Clark College.

W. S. SMALL,
Principal of Eastern High School, Washington, D. C.

E. D. STARBUCK,
University of Iowa.

L. WILLIAM STERN,
University of Breslau, Germany.

G. M. STRATTON,
University of California.

G. D. STRAYER,
Columbia University.

E. L. THORNDIKE,
Columbia University.

E. B. TITCHENER,
Cornell University.

J. B. WATSON,
Johns Hopkins University.

W. H. WINCH,
London, England.

H. K. WOLFE,
University of Nebraska.

A. H. YODER,
State Normal School, Whitewater, Wis.

VOLUME VII SEPTEMBER, 1916 NUMBER 7

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

MENTAL TESTS AND COLLEGE FRESHMEN

J. CARLETON BELL

Brooklyn Training School for Teachers

INTRODUCTION

The investigation described in this article was undertaken in the hope that the use of mental tests might be of assistance to college authorities in aiding freshmen to adjust themselves to their environment. While it is true that previous efforts of this sort have not been very successful, it was thought worth while to try a new combination of tests.

The first physical and mental measurements of freshmen were taken by Cattell (3) at Columbia College in 1896. The data accumulated in these tests were discussed by Wissler (12) in 1901, and the degree of connection between them was expressed in terms of Pearson coefficients of correlation. Wissler also took account of the class standing of these students. Table I gives some of the more pertinent correlations. The number of scores entering each correlation varied from 90 to 252.

It will be seen from the table that the correlation between marks in college subjects is fairly high (.50 to .60) but the correlations of class standing with tests or of the tests with each other are very low. This may be due to the fact that the traits measured have very little inherent connection with each other, or that the tests and the methods of giving them were such as to obscure the relationships that exist, or finally that the group is so nearly homogeneous that the variations in the test results are chiefly due to chance. It is interesting to note that the highest correla-

TABLE I

Wissler's Coefficients of Correlation

Reaction time—Cancellation of A's	—.05
Reaction time—Naming colors	.15
Reaction time—Association time	.08
Reaction time—Movement time	.14
Naming colors—Cancellation of A's	.21
Naming colors—Movement time	.19
Drawing line—Bisecting line	.38
Accuracy—Speed in cancellation of A's	—.28
Auditory memory—Visual memory (correctly written)	.29
Auditory memory—Visual memory (correctly placed)	.39
Auditory memory—Logical memory (correctly written)	.05
Auditory memory—Logical memory (correctly placed)	.04
Logical memory—Retrospective memory	—.07
Force of movement—Drawing line	—.08
Logical memory—Naming colors	.03
Logical memory—Reaction time	.12
Auditory memory—Reaction time	.17
Visual memory—Reaction time	.06
Pitch memory—Reaction time	.01
Class standing—Reaction time	—.02
Class standing—Cancellation of A's	—.09
Class standing—Association time	.08
Class standing—Naming colors	.02
Class standing—Logical memory	.19
Class standing—Auditory memory	.16
Class standing—Gymnasium	.58
Latin—Mathematics	.58
Latin—Rhetoric	.55
Latin—French	.60
Latin—German	.61
Latin—Greek	.75
Rhetoric—French	.30
Rhetoric—German	.61
Rhetoric—Mathematics	.51
Gymnasium—Mathematics	.52
Mathematics—Logical memory	.11
Latin—Logical memory	.22

TABLE II

Waugh's Coefficients of Correlation

Class standing—Quickness of association (opposites)	.54
Class standing—Speed of learning (substitution)	.24
Class standing—Range of information	.47
Class standing—Retention of ideas	.40
Class standing—Steadiness of hand	—.43
Range of information—Quickness of association	.32
Range of information—Steadiness of hand	—.20
Ingenuity (puzzle box)—Speed of learning	.51
Ingenuity (puzzle box)—Steadiness of hand	—.22

tion is between Latin and Greek (.75), and that class standing correlates more highly with logical memory (.19) than with any other test. The correlation between drawing a line equal to a given line and bisecting a given line (.38) is what we might expect, since the two attitudes are so nearly alike. The close relationship between visual and auditory memory (.39) is not surprising.

At the meeting of the American Psychological Association in Chicago, December, 1916, Waugh (11) read a paper on "A new Mental Diagnosis of College Students," a newspaper abstract of which is available. In 1912 Waugh gave the following mental tests individually to freshmen in Beloit College: (1) Concentration of attention (cancellation of A's); (2) range of information (Whipple list); (3) speed of learning (substitution tests); (4) quickness of association (opposites); (5) ingenuity (puzzle box); (6) steadiness of hand at rest; and (7) memory for a passage (immediately after reading and after a lapse of two weeks). Pearson coefficients of correlation are indicated in Table II.

All of these correlations are considerably higher than those obtained by Wissler, but whether it is on account of the difference in tests or a difference in method is impossible to say. Three years later the same tests were given to as many of the same students as remained in college. No correlations are indicated, but the statement is made that as seniors the students had improved 4.6 per cent. in concentration of attention, 4.2 per cent. in speed of learning, 2 per cent. in quickness of association, 44.7 per cent. in range of information, and 19 per cent. in class standing.

Bingham (1) gave the following nine tests individually to 200 Dartmouth freshmen: (1) Endurance of grip; (2) tapping; (3) auditory memory; (4) memory for a passage; (5) cancellation of a given digit; (6) color naming; (7) logical relations; (8) mixed relations (analogies); and (9) perception of form. No correlations are indicated, but the results are presented in "ogive" percentile graphs, so that it will be easy to compare the results obtained from other individuals with those already plotted.

Hollingworth (5), working with 13 students in the psychological laboratory, found that correlations between tests became much higher as the subjects approached the limits of improvement with practice. Table III shows the correlations at the extremes of practice.

TABLE III

Effect of Practice on Correlations

	Trial	Tapping	Coördination	Discrimination	Color Naming	Opposites	Average
Adding	1	.45	—.21	.23	.26	.23	.19
	205	.57	.16	.15	.76	.76	.48
Opposites	1	.36	—.14	—.46	.53		.10
	205	.55	.50	.39	.88		.62
Color Naming	1	.36	—.28	—.13			.15
	205	.52	.47	.41			.61
Discrimination	1	—.08	.11				—.07
	205	.59	.19				.35
Coördination	1	—.25					—.15
	205	.39					.34
Tapping	1						.17
	205						.52

In the course of a study of variations in efficiency during the day Gates (4) found that the correlations between nine mental tests with 172 pupils of grades VI, VII and VIII showed a tendency to be low. The results are seen in Table IV.

TABLE IV

Tests of School Children

	Multiplication	Addition	Speed of M'v't	Acc. of M'v't	Completion	Cancellation	Aud. Mem.	Vis. Mem.	Recognition
Multiplication		.72	.03	.04	.11	.20	—.01	—.07	.04
Addition			.01	.07	.07	.24	—.06	—.13	—.07
Speed of M'v't				—.76	.21	.29	—.06	.09	.03
Acc. of M'v't					—.07	—.17	.09	.08	.14
Completion						.15	.14	.18	.17
Cancellation							.03	.04	.09
Aud. Mem.								.57	.17
Vis. Mem.									.11
All Tests	.13	.12	—.02	—.07	.12	.11	.10	.11	.10

The only high correlations are multiplication with addition, visual with auditory memory, and a negative correlation of speed with accuracy of movement. It is noteworthy that cancellation has a considerable correlation with multiplication, addition and speed of movement, all three involving activities in some degree common to cancellation.

Starch (8) in a study of the school marks of children in grades V, VI, VII and VIII found some remarkably high correlations, as shown in Table V. In another computation, which he thinks is more reliable, he obtained even higher correlations for some of the subjects. Perhaps the correlations of school marks are higher for elementary school pupils because teachers know more about the pupils outside of their work in the special subjects than do teachers of higher classes, and therefore are governed more by general principles in their marking. This would tend to produce greater uniformity, and, therefore, higher correlations.

TABLE V

Correlations of School Marks

	Arithmetic	Language	Geography	History	Reading	Spelling
Arithmetic		.73	.74	.73	.45	.42
Language	.73		.86	.77	.80	.77
Geography	.74	.86		.81	.83	.68
History	.73	.77	.81		.67	.37
Reading	.45	.80	.83	.67		.72
Spelling	.42	.77	.68	.37	.72	

Simpson (6) chose two groups of individuals differing as widely as possible in intellectual status, and applied the same series of tests to determine what ones gave the clearest indication of this difference in mentality. His "good" group was composed of seventeen professors and advanced students in Columbia Univerity. His "poor" group contained twenty adults found in the Salvation Army Industrial Home, a Bowery mission and other refuges for human misfits. None of the "poor group" had ever held positions requiring a high grade of intelligence. His raw correlations are shown in Table VI.

These raw correlations are for both groups together, and are considerably higher than those for each group taken separately. In the one case the deviations are taken from the median of the entire thirty-seven individuals, and since the groups were chosen to represent extremes the deviations are naturally quite marked and fairly uniform. In the other case the deviations are computed from the median of each group separately, and therefore are by no means so uniform. In order to determine what the probable true correlations would be if a very large group representing a normal distribution of human minds were taken instead of two small selected groups, Simpson has worked out elaborate

TABLE VI

Simpson's Raw Correlations

	Ebbinghaus	Hard Opp.	Mem. of Words	Easy Opp.	Canc. of A's	Mem. of Pass.	Adding	Geom. Forms	Learn. Pairs	Rec. Forms	Scroll	Compl. Words	Est. Lengths	Draw. Lengths
Ebbinghaus		.98	.94	.79	.62	.91	.71	.54	.78	.88	.55	.42	.33	.25
Hard Opp.	.98		.84	.80	.64	.81	.79	.70	.73	.74	.52	.43	.26	.25
Mem. of Words	.94	.84		.62	.55	.82	.49	.56	.73	.71	.53	.40	.28	.21
Easy Opp.	.79	.80	.62		.57	.52	.68	.53	.42	.56	.45	.29	.38	.48
Canc. of A's	.62	.64	.55	.57		.55	.54	.73	.39	.51	.39	.59	.25	.22
Mem. of Pass.	.91	.81	.82	.52	.55		.53	.57	.59	.66	.54	.31	.28	.19
Adding	.71	.79	.49	.68	.54	.53		.45	.39	.47	.51	.57	.17	.25
Geom. Forms	.54	70	.56	.53	.73	.57	.45		.35	.49	.54	.56	.25	.25
Learn. Pairs	.78	.73	.73	.42	.39	.59	.39	.35		.69	.36	.29	.26	.09
Rec. Forms	.88	.74	.71	.56	.51	.66	.47	.49	.69		.44	.37	.34	.28
Scroll	.55	.52	.53	.45	.39	.54	.51	.34	.36	.44		.31	.19	.27
Compl. Words	.42	.43	.40	.29	.59	.31	.57	.56	.29	.37	.31		.21	.07
Est. Lengths	.33	.26	.28	.38	.25	.28	.17	.25	.26	.34	.19	.21		.24
Draw. Lengths	.25	.25	.21	.48	.22	.19	.25	.25	.09	.28	.27	.07	.24	

computations of "correction for attenuation" in which the highest correlations are reduced from .10 to .20 below those given in the table. Even so, however, most of the correlations remain remarkably high, and in this respect show little agreement with the results of the present study.

In his vice-presidental address before Section L (Education), American Association for the Advancement of Science, in December, 1912, Thorndike (10) took as his subject "Educational Diagnosis." He discussed Simpson's work in some detail and proposed the following eight tests for determining intellectual status: (1) Completion; (2) opposites; (3) learning pairs; (4) recognition of forms; (5) marking false statements; (6) addition; (7) following directions; and (8) selective judgment, or selecting valid reasons for a fact.

"From the combined score made by an individual in these eight tests his general intellectual ability—his capacity, that is, for science, scholarship, and the management of ideas of all sorts—could be prophesied with a surprisingly small error . . . If each of us knew all in the company well and wrote down the names in an order of general intellectual ability, and if all of these orders were combined into an order representing the impartial judgment of the total group about each of its members, this order

would be hardly any truer than the order got by using the combined scores of the tests alone. The two orders would indeed be practically identical. There is excellent reason to believe that it is literally true that the result of two hours' tests properly chosen from those already tested gives a better diagnosis of an educated adult's general intellectual ability than the result of the judgments of two teachers or friends who have observed him in the ordinary course of life each for a thousand hours."

This address furnished the starting point for the present investigation. The question was raised, "Could these tests be used in any way to throw light on the differences in intellectual capacity among the seven hundred and more freshmen at the University of Texas?" It was out of the question to devote two hours to each student individually for the application of the tests, and while it was realized that the modification of the tests for use as mass tests might rob them of all their diagnostic value, it was thought worth while to make the attempt. After a considerable amount of preliminary experimentation nine tests were selected for use. It was decided to proceed on the principle of amount of work done within a given time limit, and the time was so chosen that even the most rapid workers were scarcely able to finish the test. The attitude induced in the students, therefore, was throughout that of the speed test.

The Method of Giving the Tests

The nine test blanks were fastened together into bundles with clips, and each student was provided with a bundle, placed face down on the desk. The students were asked not to look at any of the tests until told to do so. The tests were as follows:

(1) Cancellation of triangles. (For a reproduction of this test sheet see Simpson 6, p. 113, A.) The test sheet contained 50 triangles distributed irregularly among a large number of other geometrical figures arranged in sixteen rows. At a given signal the students were instructed to turn the sheet and to cross out triangles as rapidly as possible for one minute. They were cautioned not to omit any triangles, and not to cross out any other figures. The wording of the instructions for this and all of the tests was carefully worked out beforehand, and a typewritten copy given to each examiner with the injunction to use exactly those words and no others.

(2) Addition. (See Simpson, 6, p. 117, XIa.) Time allowed, one minute.

(3) Association. (See Thorndike, 10, p. 138, fig. 1, or Simpson, 6, p. 116, fig. 2. The words and groups of lines used were slightly different from any of those represented.) This test is called "Learning Pairs" by Simpson and Thorndike. The "study sheet" contained ten groups of short, straight lines, each followed by a familiar word. At a given signal this sheet was turned and the students were given two minutes to study it in order to fix the associations between the characters and the words. They had been told beforehand that at the expiration of the time they would turn up another sheet on which they would find the same characters in a different order, and they would be expected to write after each character the word that was associated with it. Two minutes were allowed for this writing.

(4) Recognition. (See Simpson, 6, p. 113, E, F.) This is Simpson's Recognizing Forms test. The students were told that they would turn up a sheet on which there were twenty-five different forms. They would be given three minutes to study these forms, fixing them in mind so that they would be recognized anywhere. Particular attention was to be paid to the position of the forms on the page, for some of them differed from others only in the way they were placed on the page. At the expiration of the time they were to lay this sheet face down and turn up another sheet on which they would find the same forms mixed with twenty-five others. They were asked to draw a vertical line through each form recognized from the first sheet. Two minutes were allowed for the marking.

(5) Selective Judgment. (See Bonser, 2, p. 5, VA.) Ten reasons are given in answer to the question, "Why has New York become a larger city than Boston?" Three of these reasons are valid. The student was asked to place a check mark before each reason which he thought valid. Each student was allowed to finish checking the list, and as he worked the examiner wrote on the board the digits 1, 2, 3, 4, 5, 6, etc., at intervals of five seconds. As the student finished he glanced at the board and put on his paper the last digit he saw there. This digit multiplied by five gave the number of seconds spent on the test. In making up the score, however, it was found that the time had very little significance, and it was therefore ignored.

(6) Directions, No. 1. (See Woodworth and Wells, 13.) This test consisted of both the Woodworth and Wells easy lists of "Following Directions" combined on a single sheet, making forty tasks to perform. The time allowed was one minute.

(7) Directions, No. 2. (See Woodworth and Wells, 13.) This was Woodworth and Wells' hard list. The time limit was one minute.

(8) Alternatives. (See Squire, 7, p. 493.) This consisted of fifteen sentences of the following type:

Days are $\frac{\text{longer}}{\text{shorter}}$ in summer than in winter.

The students were instructed to cross out one of the two words, leaving the statement correct. Forty-five seconds were allowed.

(9) Completion. (See Terman and Childs, 9, p. 198.) This was the Terman form of the well known completion test, "The Strength of the Eagle." Five minutes were allowed.

It will be seen that the net working time for the whole series of tests was twenty minutes. Allowing time for the distribution of the test papers, the general explanation of the tests, the detailed directions for each test, the answering of questions (which were always welcomed before the test began, but positively forbidden after the signal to start had been given), and the collection of papers, the time required was usually from forty to forty-five minutes. This so distributed the efforts of the students that little fatigue was noticed. The tests were given in November, 1913, to approximately seven hundred and fifty freshmen at the University of Texas, in groups averaging a little less than twenty each.[1]

Scoring the Results

The results were first scored on thirteen large sheets of cross section paper, the scores for each student running on the same line entirely across the page. In test No. 1 (Triangles) a column was devoted to the number of triangles crossed out, another to the number of other figures wrongly crossed (positive errors), and a third to the number of triangles omitted (negative errors). The number of positive errors was very small, but the great majority

[1] The author gratefully acknowledges the assistance in giving the tests of Professors C. S. Yoakum, A. C. Ellis, L. W. Sackett, Dr. C. T. Gray and Messrs. E. D. Jennings and D. F. McCollum, graduate students. The author is also greatly indebted to Miss Lucille M. Rawlins, M. A., for valuable help in the computation of coefficients of correlation.

of students made one or more negative errors. Test No. 2 (Addition) consisted of ten problems of two columns each. A column was set aside on the score sheet for each column to be added, and errors were indicated by check marks in the appropriate columns. A vertical line indicated the last addition attempted, and in a final column was indicated the total number of columns correctly added. In test No. 3 (Association) a column was devoted to each association, and each wrong association was scored as a positive error (+), while each omission of a word (lack of association) was scored as a negative error (—). A final column showed the total number of correct associations.

In test No. 4 (Recognition) each of the fifty figures on the second or recognition sheet was given a column. If the figure was correctly marked a check mark was placed in the column. If it was marked when it should have been left blank (false recognition), a + mark was placed in the column and it was called a positive error. If it was unmarked when it should have been marked (lack of recognition), a — mark was placed in the column and it was called a negative error. If it was correctly left unmarked, the space was left blank. Final columns showed the score (the number correctly marked), the number of positive errors and the number of negative errors.

In test No. 5 (Selective Judgment) a column was given to each of the ten reasons assigned. The significant reasons were numbers 3, 6, and 8. These were checked if correctly marked. If other reasons were marked they were called positive errors and + marks were placed in the appropriate columns. If any of the three significant reasons was not marked, it was called a negative error and a — mark was placed in the proper column. Final columns indicated the number of positive errors and the number of negative errors.

In tests No. 6 (Directions No. 1), No. 7 (Directions No. 2), and No. 8 (Alternatives) a column was set aside for each task, a check mark was made in it if the task was wrongly performed, and a vertical line showed the last task attempted. Final columns showed the scores and the number of errors in each case.

In test No. 9 (Completion) no attempt was made to use the system of weightings indicated by Terman, but each completion that did not satisfy the sense of the passage was struck out and called a positive error, blanks that had not been filled were noted

and called negative errors, and the number of the last blank attempted was taken as the score of the rate of completion. The score sheet thus showed the number of completions attempted, the number of positive errors, and the number of negative errors. The author has preserved this material, and expects to make a more detailed study of the results of this test.

Reduction to Single Measures

For purposes of comparison and correlation it was desirable to reduce the results obtained from each test to a single numerical score. In doing this an effort was made to arrive at such a scheme of weighting for each test that a perfect mark, or the highest mark, as the case might be, would approximate 100, and other marks range downward from this to 0. In test No. 1 (Triangles), since there was a total of fifty triangles, two was taken as the constant multiplier for the number of triangles crossed out, and five points were deducted for each error, positive or negative. For example, if a student crossed out 35 triangles, omitted three, and crossed out one circle, his score was 70—20=50.

In test No. 2 (Addition) there were twenty columns, and on the basis of 100 each column would be worth five points. It was thought that the errors were sufficiently penalized by the reduction of the score, and the final score taken was thus five times the number of columns correctly added. In test No. 3 (Associations) both positive and negative errors were ignored, and the score taken was ten times the number of correct associations. In test No. 4 (Recognition) there were twenty-five forms to be recognized and marked. The negative errors were sufficiently penalized by their failure to increase the score, so the final score was taken as four times the number of forms correctly marked, while a positive error of false recognition was penalized by deducting three points. Thus the score of a student who had fifteen forms correct and two wrong would be 54, while that of one who had twenty forms correct and ten wrong would be 50.

In test No. 5 (Selective Judgment) each student was given 100 to start with, 15 was deducted for each positive error, and 10 for each negative error. While test No. 6 (Directions No. 1) contained forty chances, an inspection of the raw scores showed that very rarely did a student attempt as many as twenty-five in the time allowed. Therefore the final score chosen was four times

the number attempted, and since the tasks were so very simple it was thought that the deduction of ten for every error would be none too heavy a penalty. Test No. 7 (Directions No. 2) offers twenty-one chances, but no one finished in the time allowed, so the score chosen was five times the number attempted minus ten times the number of errors. Test No. 8 (Alternatives) has fifteen tasks, which makes the nearest even multiplier seven. Ten were deducted for each error. In test No. 9 (Completion) there were 107 spaces to be completed, but only three or four students finished the entire paper in the time allowed, and each of these made several errors, so it was decided to assign a score of one for each space as far as the student went, and to deduct two for each positive error and one for each negative error.

It is frankly admitted that these weightings are arbitrary, but they were arrived at after prolonged trials with different kinds of weighting, and comparison of individual scores with each other on the basis of these weightings shows that they give a just measure of the relative performance of the students in the tests. In a very few instances the scores in tests 6–9 ran slightly over 100. In that case they were arbitrarily called 100. In perhaps ten or fifteen very low scores the method of penalizing adopted made the final score a negative quantity. In these cases it was arbitrarily called 0.

University Grades

In addition to the results of the tests the university grades for each student were recorded. At the University of Texas a grade is entered on the records of the Registrar for each term's work of twelve weeks, giving three grades for the year in a full course. The system of marking in use is the letter system, A, B, C, D, E, and F, with D as the passing mark. These letters were given the following numerical values: A=90; B=80; C=70; D=60; E and F=50, and the average numerical mark for the year in each of six groups of subjects was determined. The only subject taken by all the students tested was English. For the purposes of this study the subjects taken by freshmen were divided into six groups, as follows: English, mathematics, history, science (physics, chemistry, botany, zoölogy, geology), foreign languages (Latin, Spanish, French, German), and education. Table VII indicates the percentile scores of the freshmen in the six university subjects and the nine mental tests.

TABLE VII

Percentile Scores of University Freshmen

	English	Mathematics	History	Science	For. Lang.	Education	Triangles	Addition	Association	Recognition	Select. Judg.	Dir. No. 1	Dir. No. 2	Alternatives	Completion
No. of Students	702	506	359	546	544	129	712	709	709	730	730	746	746	729	726
Lowest Score	50	50	50	50	50	50	0	0	0	0	0	4	0	2	0
10 per cent.	53	52	52	54	52	60	36	32	39	25	26	40	26	38	19
20 per cent	57	54	55	60	54	64	45	40	48	33	42	47	34	48	25
30 per cent.	61	58	60	53	58	68	51	45	55	39	48	53	41	54	31
40 per cent.	63	62	62	66	62	71	55	50	62	43	53	58	45	60	35
50 per cent.	*65*	*64*	*64*	*69*	*65*	*72*	*58*	*55*	*69*	*47*	*57*	*62*	*50*	*65*	*39*
60 per cent.	68	68	68	72	69	74	62	60	75	51	66	66	54	69	44
70 per cent.	71	71	71	74	73	76	65	66	82	56	74	70	59	74	48
80 per cent.	74	74	74	78	78	80	69	72	91	60	78	77	67	78	54
90 per cent.	80	84	80	82	85	83	76	82	100	67	86	87	75	86	60
Highest Score	90	90	90	90	90	90	95	100	100	96	100	100	100	100	94

Examination of Table VII shows that English, mathematics, and history, despite the difference in number of students, have their grades distributed in about the same fashion from lowest to highest, *i. e.*, the different percentile groups have about the same grade boundaries. Foreign languages also have practically the same distribution below the median, but their better-class students receive relatively higher grades than those in the other groups, although it is noteworthy that the upper ten per cent. of students in mathematics are better favored than those in either English or history. The sciences start their grading higher than any of the subjects already considered, and maintain their lead as far up as the eighty percentile group, where they are overtaken by foreign languages. Education shows a much higher appreciation of its low grade students than any of the other subjects, maintaining a striking lead as far as the eighty percentile. Its better students, however, do not get as high marks as those in mathematics or foreign languages.

Turning our attention now to the tests, we are impressed at once by the strikingly high scores in Association, and the equally low ones in Completion. The Association test was obviously the easiest of all the tests. The lowest ten per cent. did better with it than any other test except Directions No. 1, and the other percentile groups increased their lead until in the best ten per cent. all

made a perfect score. Precisely seventy-one out of a total of 709 made a score of 100. In the Completion test, on the other hand, the scores start low, the median is eight points lower than in any other test, and the upper percentiles are relatively low. This was due partly to the weighting employed, but also to the fact that the test itself was found distinctly difficult. Another difficult test, the scores of which are only slightly higher than those of the Completion test, is the Recognition test. The most evenly balanced of all the tests so far as the percentile groups are concerned is Directions, No. 2. The median score is just 50, and the percentile groups are distributed with remarkable regularity below and above it.

Table VIII presents the Pearson coefficients of correlation of university marks with each other, of marks with test scores, and of test scores with each other. These were determined by the formula

$$r = \frac{\Sigma xy}{\sqrt{\Sigma x^2}\ \sqrt{\Sigma y^2}}$$

It will be seen that the correlations between university marks are fairly high, ranging from mathematics-history=.34 to English-history and science-history=.59. Why marks in science should correlate so highly with those in history is not clear. Foreign languages have the highest average correlation with other subjects (.542) and mathematics the lowest (.484). The correlations of university marks with test scores are uniformly low, the highest being English-Completion=.31. Triangles and Selective Judgment have practically no correlation with any university marks, and the other tests show a considerable variation. The relatively high correlation of English with the two Directions and the Alternatives tests is interesting, as is also that between Addition and mathematics, science and foreign languages.

Considering the relations of the test scores to each other, one is immediately impressed by the connections between the two sets of Directions, the Alternatives and the Completion test. This is the only group of relatively high correlations that exists among the tests, if we except the Association-Recognition relationship (.30) in which there is a high degree of identity in procedure. Perhaps the connection between these four tests is due to the linguistic factor which is present in all of them. The ability to

TABLE VIII

Coefficients of Correlation, University Freshmen

	English	Mathematics	History	Science	For. Lang.	Education	Triangles	Addition	Association	Recognition	Select. Judg.	Dir. No. 1	Dir. No. 2	Alternatives	Completion
English		.46	.59	.45	.58	.44	.01	.11	.20	.17	.00	.26	.27	.24	.31
Mathematics	.46		.34	.57	.52	.53	.13	.23	.16	.14	.07	.13	.12	.18	.22
History	.59	.34		.59	.56	.43	.00	.13	.12	.18	.06	.19	.06	.15	.16
Science	.45	.57	.59		.49	.48	.04	.25	.12	.19	.06	.12	.15	.13	.17
For. Lang.	.58	.52	.56	.49		.56	−.03	.25	.15	.14	.06	.19	.18	.11	.20
Education	.44	.53	.43	.48	.56		.02	.12	.02	−.08	.04	.03	.04	.25	.16
Triangles	.01	.13	.00	.04	−.03	.02		.10	.13	.08	.02	.12	.10	.21	.14
Addition	.11	.23	.13	.25	.25	.12	.10		.07	.06	.09	.14	.15	.20	.12
Association	.20	.16	.12	.12	.15	.02	.13	.07		.30	.05	.14	.11	.14	.21
Recognition	.17	.14	.18	.19	.14	−.08	.08	.06	.30		.03	.17	.17	.13	.15
Select Judg.	.00	.07	.06	.06	.06	.04	.02	.09	.05	.03		.01	−.04	.06	.09
Dir. No. 1	.26	.13	.19	.12	.19	.03	.12	.14	.14	.17	.01		.33	.34	.36
Dir. No. 2	.27	.12	.06	.15	.18	.04	.10	.15	.11	.17	−.04	.33		.23	.30
Alternatives	.24	.18	.15	.13	.11	.25	.21	.20	.14	.13	.06	.34	.23		.29
Completion	.31	.22	.16	.17	.20	.16	.14	.12	.21	.15	.09	.36	.30	.29	

quickly sense the meaning of the printed words and perform the simple task that is required is the thing that contributes to a high score in any one of these tests. In so far then we would seem justified in assuming a more or less general "linguistic sense" or skill in dealing with linguistic problems, which is involved in all four of these tests, and which also contributes largely to success in college classes.

On the other hand the Selective Judgment and Triangles tests have practically no correlation with any of the tests or college subjects. If we compute the average correlation of each test with all the others and arrange them in descending order we have the following array: (1) Completion, .206; (2) Alternatives, .190; (3) Directions, No. 1, .181; (4) Directions, No. 2, .155; (5) Addition, .144; (6) Association, .137; (7) Recognition, .131; (8) Triangles, .076; (9) Selective Judgment, .043.

Scores of Best and Poorest Students

In a study of the individual scores for contrasts in results it was noted that thirty-seven students had an average class standing of 80 or higher, and that precisely the same number had an average class standing of 52 or lower. The individual scores for these two

TABLE IX

Scores of Best Students

	Av. Class Standing	Triangles	Addition	Association	Recognition	Select. Judg.	Dir. No. 1	Dir. No. 2	Alternatives	Completion	Total
Medians		58	55	69	47	57	62	50	65	39	502
1. Litsey, Mary	88	71	55	100	44	100	100	75	78	69	692
2. Davis, M.	88	60	60	90	75	80	92	60	77	49	643
3. Masters, Mildred	88	65	80	70	57	25	100	75	61	76	609
4. Trice, L.	88	62	60	100	53	85	92	55	49	59	605
5. Ware, L.	88	66	55	60	39	55	70	65	78	42	530
6. August, Maxie	87	64	95	60	27	40	72	55	53	40	506
7. Lewin, B. D.	86	65	80	50	66	80	100	75	88	89	693
8. Shelton, E.	85	76	80	70	36	100	18	65	100	92	637
9. Wynn, F.	85	64	45	100	85	55	76	45	60	73	603
10. Riley, M.	85	46	40	70	40	75	68	55	70	50	514
11. Wood, O. J.	85	45	100	60	41	30	62	35	65	0	438
12. Sweet, M.	84	76	55	80	96	55	58	85	77	93	675
13. Osborne, E.	84	64	65	100	84	40	66	75	81	57	632
14. Higdon, J. E.	84	66	55	50	66	85	100	35	85	48	590
15. Earhart, J.	84	64	65	80	56	40	54	55	36	52	502
16. Beckmann, W.	84	50	45	70	35	55	52	60	67	44	478
17. Wilmeth, R. D.	83	60	45	100	65	40	62	55	60	41	528
18. Girard, F.	83	0	70	60	26	50	75	55	50	50	437
19. Nagle, R.	82	53	55	90	73	70	94	40	78	66	619
20. Baker, E.	82	47	75	80	59	100	80	40	60	54	595
21. Snyder, Roy	82	62	70	80	71	40	40	50	36	55	504
22. Warren, W.	82	22	65	50	46	55	56	50	39	25	408
23. Rawalt, N.	81	72	60	100	64	55	100	50	98	48	647
24. Bagnall, B.	81	55	50	100	73	70	80	75	100	26	629
25. Blair, W. C.	81	69	80	100	52	55	64	65	50	56	591
26. Harrold, C.	81	72	40	100	63	55	46	65	91	43	575
27. Lorelare, M.	81	32	70	80	56	50	72	40	50	55	505
28. Brand, C. E.	81	38	55	80	44	55	58	45	47	48	470
29. Burt, Helen	81	55	50	40	26	5	40	45	88	37	386
30. Fish, J. W.	80	76	75	100	57	85	56	55	71	50	625
31. Atkinson, J. B.	80	57	50	100	80	70	76	55	44	57	589
32. Lockett, H. S.	80	45	70	60	64	45	72	70	63	55	544
33. Cunningham, C.	80	62	55	60	58	55	80	60	71	42	543
34. Nelson, M.	80	54	75	70	38	20	96	60	48	55	516
35. Mobley, Helen	80	60	45	60	39	85	76	25	84	20	494
36. Shelmire, B.	80	43	40	70	56	70	54	10	64	54	461
37. Baker, Lois	80	0	45	60	40	70	52	45	74	32	418
Number of scores above median		22	19	25	23	15	23	23	19	31	27

TABLE X

Scores of Poorest Students

	Av. Class Standing	Triangles	Addition	Association	Recognition	Select. Judg.	Dir. No. 1.	Dir. No. 2	Alternatives	Completion	Total
Medians		58	55	69	47	57	62	50	65	39	502
1. Hopkins, H. R.	52	76	60	100	32	85	64	50	77	48	592
2. Givens, M.	52	52	90	70	54	85	60	35	40	44	530
3. McGregor, R.	52	59	60	50	50	55	84	50	77	45	530
4. Crow, F. A.	52	72	65	20	64	55	56	75	67	44	518
5. Skiles, S. A.	52	73	65	80	32	40	72	60	47	33	502
6. Christoffer, O. T.	52	76	75	40	49	45	38	40	84	48	494
7. Chambers, F. E.	52	77	50	69	41	30	100	10	70	23	470
8. Cleveland, C R	52	62	30	30	43	85	56	65	60	29	460
9. McKnight, C. B.	52	80	50	60	42	40	66	50	22	45	455
10. Gaudy, W. S.	52	58	55	70	41	70	38	30	50	31	443
11. Davidson, J. T.	52	0	50	90	36	85	56	10	60	50	437
12. Martin, Floss	52	60	30	40	59	75	44	45	32	49	434
13. Dickey, W.	52	48	30	20	49	55	56	55	70	51	434
14. Clifton, Carrie	52	56	55	50	38	70	80	10	70	0	429
15. Lang, M. A.	52	78	50	70	47	10	50	40	60	13	418
16. Simpson, S.	52	64	30	50	50	60	34	40	30	32	390
17. Porcher, S.	52	56	35	20	48	40	44	40	70	29	382
18. Dyer, Earl	52	22	35	80	34	40	48	15	53	48	375
19. Harlen, H. L.	52	66	50	70	39	10	26	25	60	27	373
20. Sutherland, G.	52	18	55	20	41	55	80	45	40	12	366
21. Coats, R. H.	52	38	30	10	55	25	44	25	56	5	288
22. Williams, Horace	51	0	25	40	42	55	84	30	70	28	374
23. Soule, D. F.	51	58	25	69	19	40	16	15	16	31	289
24. McCambell, C. A.	50	41	80	50	40	85	70	40	70	40	516
25. Creaton, L. B.	50	60	35	100	46	85	56	40	63	27	512
26. Terhune, S.	50	80	65	100	24	30	72	45	49	26	491
27. Lindsey, George	50	31	45	70	62	70	46	50	53	59	486
28. June, E.	50	61	70	80	38	25	56	20	64	50	464
29. Floyd, J.	50	63	90	40	63	45	44	55	30	28	458
30. Maxon, W.	50	86	25	70	57	40	44	55	50	29	456
31. Stiles, Eudora	50	13	60	90	24	70	50	60	56	33	456
32. Graves, L.	50	62	60	30	40	55	48	70	47	29	441
33. Thompson, P.	50	64	25	70	44	70	44	35	47	25	424
34. Farley, R. H.	50	16	55	80	35	55	52	35	43	26	397
35. Hutchins, W. M.	50	60	40	40	66	45	28	30	33	33	375
36. Marshall, L. B.	50	60	35	0	40	70	60	35	36	32	368
37. Heard, Thos.	50	65	30	40	55	0	48	35	22	45	340
Number of scores below median		13	21	19	22	23	27	25	27	23	30

groups of students are found in Tables IX and X. The average of the total scores for the group of the best students was 563, or 61 points above the median for all students. The highest was 693, the lowest 386. For the group of poorest students the average was 438, or 64 points below the median, the highest was 592, and the lowest 288. In the best group the number of scores above the median is especially high in the Completion test, and is fairly high in the totals, the Association, Recognition, both Directions, and the Triangles tests. In the poorest group the number of scores below the median is striking in the totals, and in the Alternatives, the two Directions, the Completion, and the Selective Judgment tests. It is possible that the chief element of difference here, as in the university work, is lack of earnest and persistent application to the tasks set. But the fact that many of the poorest group do well in some tests, while many of the best group make low scores in some tests, would indicate that there are other factors involved.

Conclusions

1. The correlations between freshmen university grades vary from .34 (mathematics-history) to .59 (English-history, science-history).
2. The highest correlation between class marks and test scores is .31 (English-Completion).
3. Among the tests themselves the highest correlations are found between the Association and Recognition tests, and between the Directions, Alternatives and Completion tests.
4. There is a considerable difference in the results of the tests with the best and the poorest students, but the scores are so variable as to be of little value for individual diagnosis. This conclusion does not in any way invalidate Professor Thorndike's claims for the tests he proposed, for those tests were in part different, were to be given individually, were to have no time limit, and were to be given in four different forms to each individual.
5. There is little doubt that if a larger number of trials with similar forms of each test had been possible, the scores would have become "steadier," and the correlations between the tests higher.

BIBLIOGRAPHY

1. BINGHAM, W. V. *Some Norms of Dartmouth Freshmen.* Journal of Educational Psychology, 7: 1916, 129–142.

2. BONSER, FREDERICK G. *The Reasoning Ability of Children of the Fourth, Fifth and Sixth School Grades.* Teachers College, Columbia University Contributions to Education, No. 37, 1910. Pp. vii, 133.

3. CATTELL, J. MCKEEN, AND FARRAND, LIVINGSTON. *Physical and Mental Measurements of the Students of Columbia University.* Psychological Review, 3: 1896, 618–648.

4. GATES, ARTHUR I. *Variations in Efficiency during the Day, together with Practice Effects, Sex Differences, and Correlations.* University of California Publications in Psychology, 2: No. 1, March 16, 1916. Pp. 156.

5. HOLLINGWORTH, H. L. *Correlations of Abilities as Affected by Practice.* Journal of Educational Psychology, 4: 1913, 405–414.

6. SIMPSON, BENJAMIN R. *Correlations of Mental Abilities.* Teachers College, Columbia University Contributions to Education, No. 53, 1912. Pp. 122.

7. SQUIRE, CARRIE R. *Graded Mental Tests.* Journal of Educational Psychology, 3: 1912, 363–380, 430–443, 493–506.

8. STARCH, DANIEL. *Correlations among Abilities in School Studies* Journal of Educational Psychology, 4: 1913, 415–418.

9. TERMAN, LEWIS M., AND CHILDS, H. G. *A Tentative Revision and Extension of the Binet-Simon Measuring Scale of Intelligence.* Journal of Educational Psychology, 3: 1912, 61–74, 133–143, 198–208, 277–289.

10. THORNDIKE, E. L. *Educational Diagnosis.* Science, N. S. 37: No. 943, Jan. 24, 1913. 133–142.

11. WAUGH, KARL T. *A New Mental Diagnosis of the College Student.* The New York Times, Magazine Supplement, Jan. 2, 1916. 12–13.

12. WISSLER, CLARK. *The Correlation of Mental and Physical Tests.* Psychological Review Monograph Supplements, 3: 1901, No. 6, 1–61.

13. WOODWORTH, R. S., AND WELLS, F. L. *Association Tests.* Psychological Monographs, Vol. 13, No. 57, 1911. Pp. 1–85.

THE STANDARDIZATION OF CERTAIN MENTAL TESTS FOR TEN-YEAR-OLD CHILDREN[1]

HOMER W. ANDERSON AND GEO. H. HILLIARD

Edited and directed by Professor Irving King, University of Iowa.

Problem

The aim of the present study is to standardize as far as possible certain mental tests for ten-year-old children. The tests used involve the following mental processes:

1. Rate of perception,—A tests
2. Immediate Memory,—Whipple's 13 picture chart
3. Association,—Two opposites tests (Woodworth and Wells)
4. Learning,—Association of numbers with certain geometrical forms
5. Linguistic invention,—Sentence, boy river ball
6. Mechanical invention,—2 rectangle triangle tests
7. Judgment
   (a) Selective,—Reasons for grass growing in summer
   (b) Problem questions

Material

These tests were given to 115 unselected ten-year-old children in the Iowa City Public Schools. There were 51 boys and 64 girls. The Grade distribution was as follows:

	Grades						
	I	II	III	IV	V	VI	Total
Boys.....		5	9	15	17	5	51
Girls.....1		1	5	18	31	8	64
Total.....1		6	14	33	48	13	115

Description and Method of Giving the Test

In General. The two co-workers in these tests tested about the same number of pupils. Each pupil was given the tests individually, removed as far as possible from all disturbing elements. It was the aim to secure as nearly as possible uniform conditions. This was secured by means of separate rooms, and secluded por-

[1] A study from the Educational Psychology Seminar, 1914–1915.

tions of the hall. In all cases the subject was placed at a table directly opposite the experimenter, and was at all times made to feel at perfect ease as far as possible. The stop watch was used in taking the time. Both experimenters used the same methods of approach to the tests and adopted standard methods of giving each of the tests. These will be described in the succeeding sections.

Rate of Perception—This test was a combination of two A tests. "A test" number one, consisting of 50 capital A's distributed irregularly on a sheet of paper, was given first to each child. The subject was told to draw a line through each A as quickly as he could. Time was taken from the instant the first check was made and ended with the last. The time was recorded on the score card. For the final score in this test a penalty of two points (seconds) for each A omitted was added to the time scored.

The second part of the test was A test No. 2, in which the A's are intermingled with other letters but at practically the same intervals apart as in test No. 1. The child was told to pick out the A's as fast as he could and mark them as in the preceding test. The time was taken in the same manner as described above. A penalty of six points (seconds) was added for each omission.

In order to secure as nearly as possible the rate of perception of the child the attempt was made to eliminate the motor factor involved in a test of this kind. This was done by subtracting the score of each pupil in test No. 1 from his score in test No. 2. Test No. 1 is almost entirely a test of motor ability and represents practically the same motor adjustments as test No. 2. The above process of obtaining the child's perceptive ability was suggested and used by Squire[2] and seems to be a more important index of ability than the ordinary A test. The objection may be raised that this method does not eliminate the eye-motor factor. This is true, but it seems a negligible factor for the purpose of this test.

Immediate Memory—The ability in immediate memory was tested by means of Whipple's chart of thirteen pictures of familiar objects. The child was directed as follows: The card was placed face down in front of the child so that it could be easily exposed. The subject was given these instructions: "I have here a card with some pictures on it. I want you to look at it carefully for a

[2] SQUIRE, CARRIE R. *Graded Mental Tests.* Journal of Educational Psychology, 3: 1912, p. 370.

little while. I am then going to put it down so you can't see it and ask you to tell me what you saw." The card was then exposed for fifteen seconds, when it was placed out of sight. After an interval of one minute the child was asked to tell what was seen.

The score given in this test was simply the number of pictures remembered. If eight were recalled, eight was the score obtained, etc.

Association—Two tests of opposites of twenty words each were used to test controlled association.[3] Test No. 1 was given first. The list was placed face down before the child. It was first ascertained whether the child knew what is meant by an opposite of a word. The procedure was as follows: "I have here a list of words and I want you to give the opposite of each word. Do you know what the opposite of a word is?" No matter what the answer of the child the observer continued. "If a knife is not dull what is it?" If the child answered "sharp" without hesitation the test was then given. If not, he was told the answer and given a second example. "If a person is not young, what is he?" If the child still seemed uncertain, he was worked with further until the experimenter was satisfied that he knew what was wanted. The subject was then given these instructions: "You know, now, what the opposite of a word is. I want you to give the opposites of these words as fast as you can. Give just the opposite." Time was taken from the moment the list of words was exposed until the last word had been given. On a separate list the experimenter checked all omissions, errors, and those partly right.

The final score was obtained by adding to the time six seconds for each omission or error and three for each opposite partly right.

Learning[4]—The card containing the five numbered geometrical figures was placed face down before the child. The following instructions were given: "I have here a card with some triangles, squares, diamonds and circles. Each one of these squares, triangles and circles is numbered. I want you to look at them a while so that you will know the number of each square, triangle,

[3] For these lists see *Association Tests*, Psychological Review. Monograph Supplements, Vol. XIII, No. 5. 1911.

[4] A sample of this test will be sent to anyone desiring to see it.

diamond and circle. Now study carefully." The card was then exposed ten seconds. A card containing fourteen blank geometrical figures was then handed to the child, who was asked to put in the right number in each square, triangle, circle, etc. If no mistakes were made on this card the test was considered finished. If any mistakes did occur, the child was told that all were not right. He was given five seconds more in which to study the numbered figures. A second blank was handed him. This sheet also contained fourteen figures, but in different order from the first blank. If mistakes still occurred he was given another study period of five seconds until four such had been given. If he did not succeed the fourth time the test was ended.

The score was derived by means of a system of arbitrary weighting. If no errors occurred after one learning the score received was 0. But if errors occurred after the first learning, one point penalty was given for each error out of the fourteen chances. For the second learning, two points penalty was given for each error; third learning, three points; and fourth learning, four points. The final score was the sum of all the penalties.

Linguistic Invention—This test was based upon the rapidity with which the child was able to give a perfect sentence containing the words "boy, river and ball." The experimenter first tested the child's knowledge of a sentence. Certain of this point, he proceeded with the test as follows: "I want you to make one sentence with those three words in it. Put all the words in the same sentence. These are the words; boy, river and ball." Time was recorded from the instant the last word, "ball," was spoken, until the child had given a sentence as requested.

The score was entirely one of time.

Mechanical Invention—In order to determine this ability two tests were given. Test No. 1 consisted in placing the two triangles together so as to form a rectangle.[5] A rectangular piece of cardboard and at the same time the two triangles, placed in such a position that the solution of the problem was not thereby suggested, were placed upon the table before the child. Directions given to the subject were to put the triangles together as fast as he could, so that they would look like the rectangle. Time was recorded from that instant until the completion of the problem. A time limit of three minutes was given. Any one failing in this time was scored "Failed."

[5] Binet test No. 55 of the 1908 series.

The second test was administered in exactly the same way. This consisted of three triangles formed by dividing one of the triangular cards of the previous test by cutting from the corner opposite the hypotenuse to the hypotenuse. This division was not perpendicular to the hypotenuse.

The score in both of these tests was the time taken to perform them, or "Failed" if not completed in three minutes.

Judgment—

1. Selective Judgment: This test is composed of ten reasons why grass grows in summer and not in winter. The score in this test was taken to consist of the number of reasons correctly evaluated.

Judgment Test

The following reasons have been given to show why grass grows in summer and not in winter. Write *yes* after those reasons you consider good and *no* after the reasons you think poor.

1. Summer is warm and winter is cold.
2. Grass is green.
3. Grass needs warmth.
4. Grass needs sunshine.
5. Cows and horses eat grass.
6. Grass needs moisture.
7. The ground is frozen in winter.
8. Children skate in winter.
9. Grass is sometimes cut for hay.
10. Grass could not grow in the frozen north.

The score in this test was taken to consist of the number of reasons correctly evaluated.

2. Problem Questions: This test consisted of the following five problem questions.

(1) A boy said, "I know ten good men who are doctors, and ten bad men who are policemen. So doctors are better than policemen." Did he prove it? Why? Or why not?

(2) If all the boys who are good in arithmetic are good in spelling, will all the boys who are good in spelling be good in arithmetic? Why? Why not?

(3) If there were no bread or flour would everyone starve?

(4) Is this true? "The more we eat, the more we grow."

(5) If there were no schools, would children learn anything?

These were asked orally by the experimenter and answered orally.

The answers were scored as follows: If correct and reasoned correctly a credit of *two* was given; if correct but not correctly reasoned, that is, answered intuitively, one credit was given. The final score was the total of the scores for each problem.

Results

1. *The rate of perception test.* The median scores of the boys and girls are given separately by grades in Figure 1. It should be remembered that the lower the score number the better it is. It will be seen that the ten-year-old boys and girls who are in the fifth and sixth grades are closely alike, while in all the grades below, the boys are considerably inferior to the girls. In grades four, five and six the scores range from slightly less than 100 to over 250; in the grades below the fourth the range is from less than 150 to 300. The median score for all ten-year-old boys is 195 and for all ten-year-old girls 161.

One might conclude from the grade correlation shown in this figure that the test as here given is a good measure of one phase of the mental ability of children of this age. We believe that the modification of the "A test" here used (suggested by Squire) has distinct advantages over the ordinary type of "A test." The test as usually given never shows a high correlation with mental ability, but in this modification the correlation is distinct with grades and it may be assumed, fairly, we think, that the progress in the grades is at least some measure of ability.

2. *Immediate memory.* In every grade there was great variability of these ten-year-olds in the scores on immediate memory. In the sixth grade the range for boys is 6 to 9 points. For the girls of this grade the range is greater, 4 to 10. The variations in the median accomplishment of boys and girls by grades is shown in Figure 2. From this it will be seen that there is little or no correlation with grade advancement. The results here obtained suggest the need of further investigation of this and similar tests. Other tests in immediate memory in school children have led to opposite conclusions from those presented here, but in many of these tests not as many children were included as in our experiment. We suggest, supposing that our results are accurate, that it is not immediate memory but retentive memory that is correlated with school progress.

3. *Association.* The same wide range of abilities was displayed in this test as in the preceding. Both boys and girls in each grade show the same variations, from the best score, *forty*, to the poorest, *two hundred.* Figures 3 and 4 give the variations of the medians by sex and grade in these two opposites

tests. These curves show that list No. 1 is easier than list No. 2.

Sex variations. In opposites test No. 1 the medians of the boys are better than those of the girls in all grades but the three lowest. In the other list, No. 2, the superiority of the boys is quite as marked. (See Figures 3 and 4.) The medians of the boys in the two tests are 69 and 92 and those of the girls are 84 and 105. Bonser[6] found in the groups he tested that girls in the 5th and 6th grades excelled the boys in the opposites test. He also found much the same grade correlations as appear in our study. Practically all who have used opposites tests have found good correlations with other measures of ability.

4. *Learning.* Wide variations of ability appear in this test. Twenty per cent. of these ten-year-olds were able to accomplish the test without error in the first trial, *i. e.* after one learning, 22 per cent. required two trials or two learnings, 22 per cent. required three trials, 8 per cent. four trials and 28 per cent. failed altogether. In other words 42 per cent. were able to perform the test in one or two learnings and 58 per cent. were not.

Sex differences. The girls' median scores excel those of the boys in all grades except the 4th and 5th. But on the whole the girls have only a slight lead over the boys, their median being 16, in terms of penalties, and that of the boys being 17. The sex differences by grades are shown in Figure 5. It will be seen that the grade correlations are good except in the case of the lower grade girls and as there were only 7 in this group the median score is not of much significance. It may be, however, that this group of seven girls is a little above the average for in the "perception test" they also proved to be superior to the girls of the 4th grade.

5. *Linguistic invention.* Sex variations are not striking but the girls are slightly superior to the boys, the median time of the former being 8.3 seconds and of the latter 10.2 seconds. The correlation with grade advancement is good. (See Figure 6.) Boys and girls together made the following scores by grades: Grades 1st to 3rd, 17 seconds, 4th grade, 9.4 seconds, 5th grade 8.3 seconds, and 6th grade 5 seconds.

6. *Mechanical invention, No. 1* (to reconstruct a diagonally bisected rectangular card). Forty of our ten-year-olds were able to perform the test in five seconds or less, and fifty-two

[6] See his "*Reasoning Abilities of School Children.*"

were able to do it in ten seconds or less, 50 per cent. completed the test in fourteen seconds or less and 25 per cent. were not able to do it in less than 47.3 seconds. The middle 50 per cent. lies between 3.4 and 47.3 seconds. The sex and grade differences in medians are not shown here for want of space; however the boys as a whole led the girls by a median score of 9 to 17 seconds. The grade correlation was not perfect for the girls but the ability of the boys varies throughout directly with the grade.

7. *Mechanical invention, No. 2.* This test is obviously more difficult than No. 1. Of all the ten-year-olds, 9 per cent. finished the test in less than one minute, 14 per cent. took from one to two minutes, 9 per cent. from 2 to 3 minutes and 68 per cent. failed in the 3 minutes allotted to it. It would seem therefore, that the test as here given is too hard for ten-year-olds. Squire does not agree with this conclusion. She found that the 9 ten-year-old children in her unretarded group completed the test with an average score of 89.3 seconds. In our test, the 35 children who were able to complete the test at all made an average score of 90 seconds, and the remainder of the group, 75 in all, were not able to do it all in three minutes. Squire also found that her ten-year-old group completed test No. 1 in the average time of 2.1 seconds. We found that 110 ten-year-olds completed the test in a median time of 14 seconds, approximately the average time. It does not seem that Squire's results are valid for a large group of non-selected ten-year-olds.

The sex and grade differences in the medians for this test are given in Figure 7. The grade correlation is fairly good. The girls are slightly better than the boys on the whole, 34 per cent. of them finishing in 3 minutes or less and 30 per cent. of the boys succeeded in 3 minutes or less.

8. *Judgment, selective.* Reasons why grass grows in summer. The same wide variation in ability is revealed by this as by the preceding tests. The median score for these 110 children was 4.7 and the middle 50 per cent. of the group lies between the scores of 4.2 and 5.9.

Sex and grade differences. Figure 8 shows that there was little difference between the sexes. The grade correlation is not so pronounced as in the previous tests. Judging from the high median score of the 6th grade as compared with the other grades, it is possible that the test does differentiate the brighter from the normal and retarded groups. It was observed that, with very few exceptions, the child in the 5th grade and lower

answered each question simply with reference to its intrinsic validity and lost sight completely of the problem in hand. For instance the statement "Grass is green" would be checked as a correct statement, the child failing to see that the question was to select the good reasons for grass growing in summer. But with the ten-year-olds who were in the 6th grade this was not the case. They alone seemed to appreciate the problem. We suggest that this test is chiefly valuable for this age as a means of differentiating the superior from the average child. A score of 7 rights or even of 6 would be an indication of superiority.

9. *Judgment, problem questions.* The median score for our group of 115 ten-year-olds was 5.4. The middle 50 per cent. is distributed between 2.8 and 8.3.

Sex differences. With the exception of grade 6 the differences between the boys and girls are negligible. This is brought out in Figure 9, The median for the entire group of girls is 5.9 and for the boys, 5.1. This same figure also shows how far the median standings of each grade are related to grade progress.

A Comparison of the Aggregate Achievement of Boys and Girls

In order to make these comparisons both sexes were ranked in the same series. This was based upon the total of the rankings in all the tests. The one with the lowest total would then rank 1, with next to the lowest total, 2, etc. A quartile distribution was then made and it was found that 14 or 28 per cent. of the boys were in the upper quartile as against 13 or 22 per cent. of the girls. This is shown in Table 1. It appears, also, that 64 per cent. of the boys were in the upper 50 per cent. and only 37 per cent. of the girls. From these figures the only conclusion is this—Boys excel in the performance of these tests.

TABLE I.

Showing the Quartile Distribution of the Rankings of Boys and Girls.

	Boys		Girls	
	No.	%	No.	%
1st 25%	14	28	13	22
2nd 25%	18	36	9	15
3rd 25%	10	20	17	29
4th 25%	8	16	20	34
Upper 50%	32	64	22	37
Lower 50%	18	36	37	13

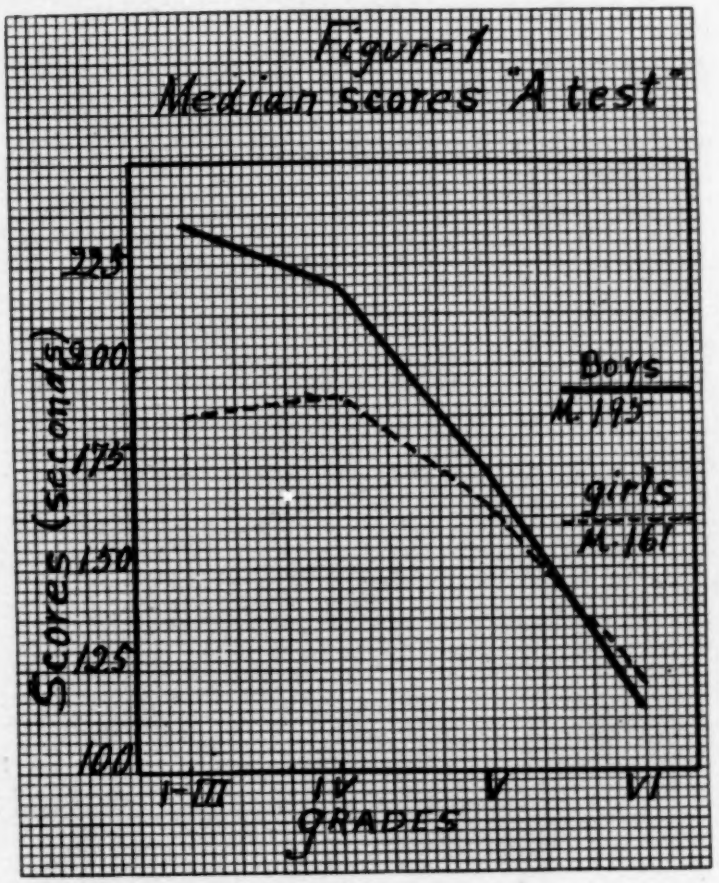
Figure 1
Median scores "A test"
225
200
175
150
125
100
Scores (seconds)
Boys
M. 195
girls
M. 161
I-III
IV
V
VI
GRADES

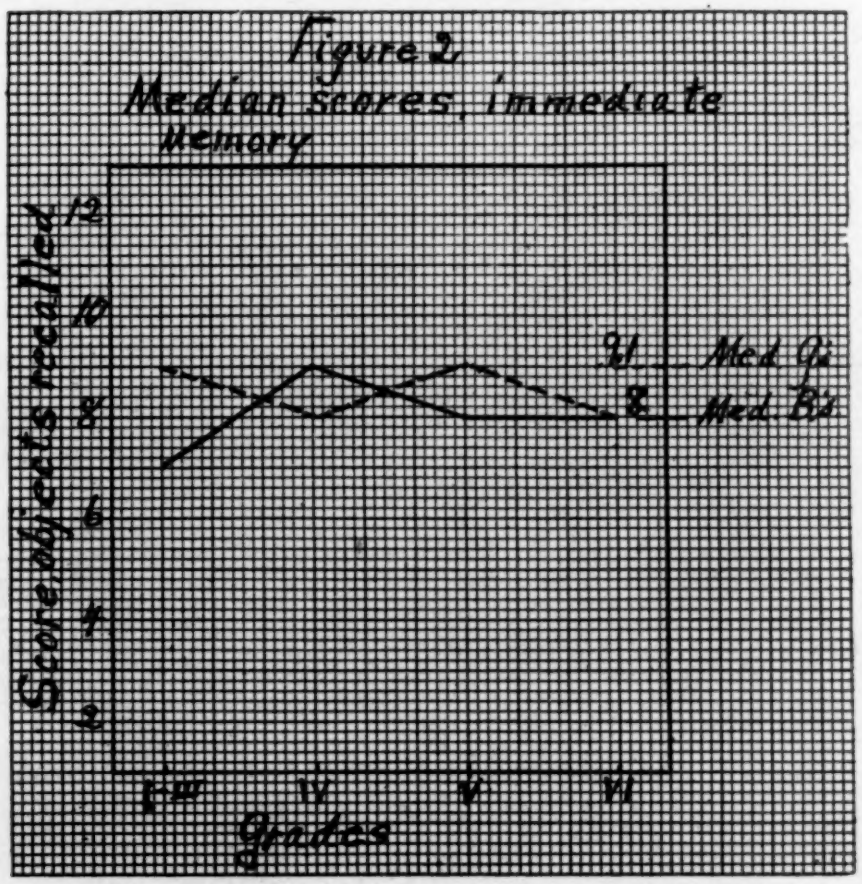
Figure 2
Median scores, immediate
Memory
Scores, ideas recalled
12
10
8
6
4
2
91
Med Q2
Med Bis
IV
V
VI
Grades

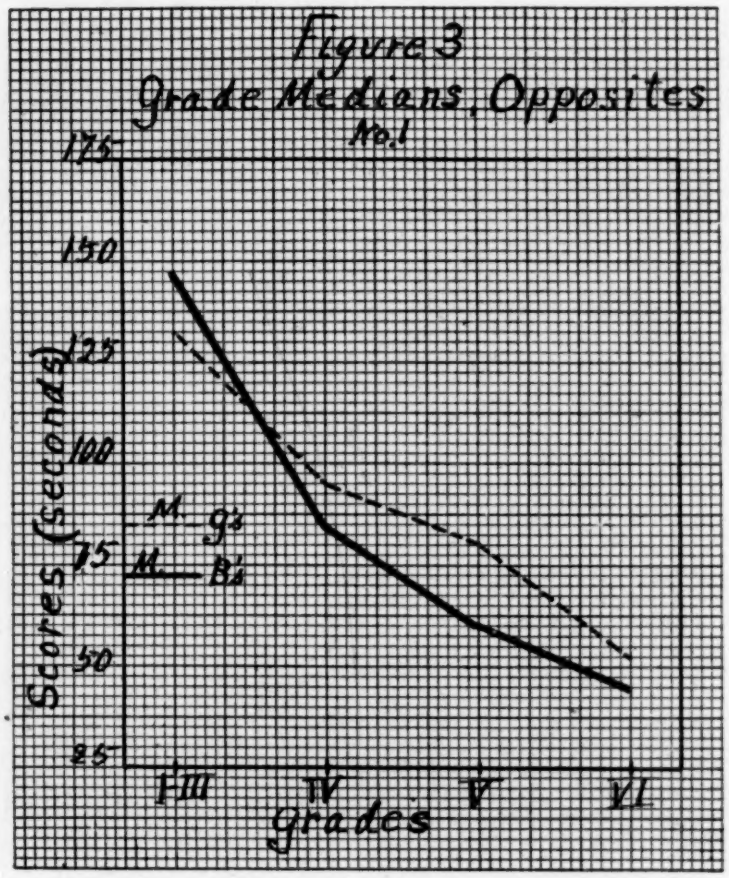
Figure 3
Grade Medians, Opposites
No. 1
175
150
125
100
75
50
25
Scores (seconds)
M. Gs
M. Bs
III
IV
V
VI
Grades

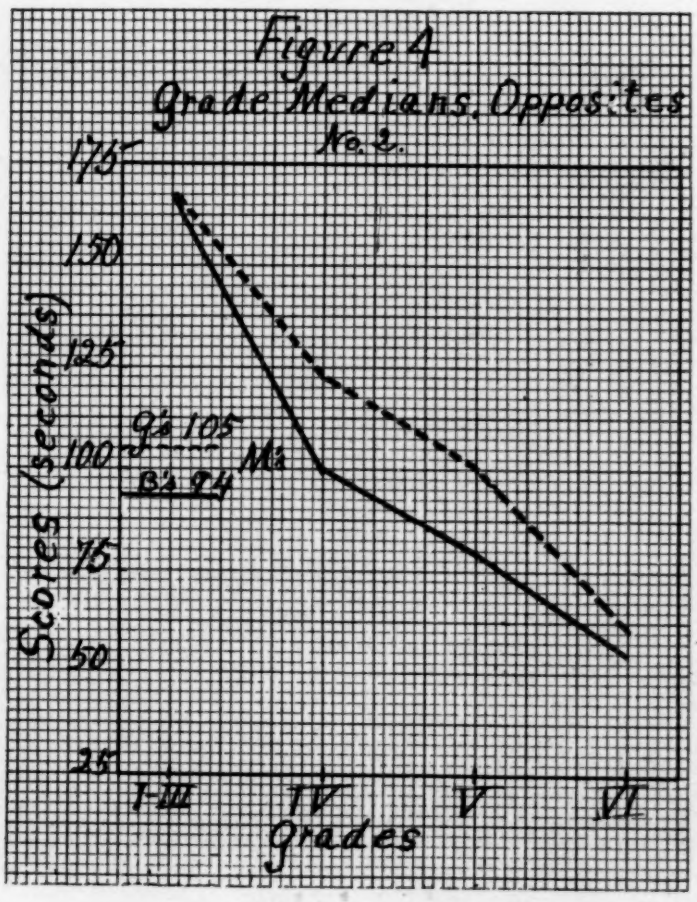
Figure 4
Grade Medians, Opposites
No. 2.
175
150
125
100
75
50
25
Scores (seconds)
Gs 105
Ms
Bs 94
I-III
IV
V
VI
Grades

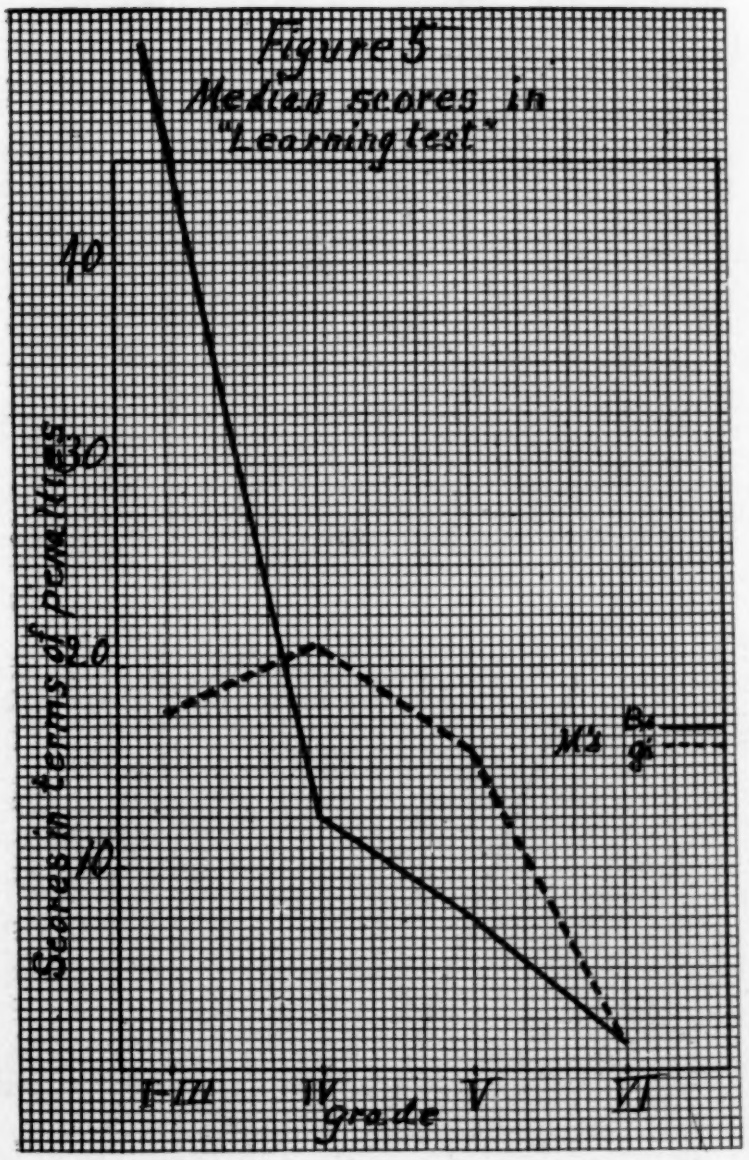
Figure 5
Median scores in
"Learning test"
10
30
20
10
Scores in terms of
I-III
IV grade
V
VI

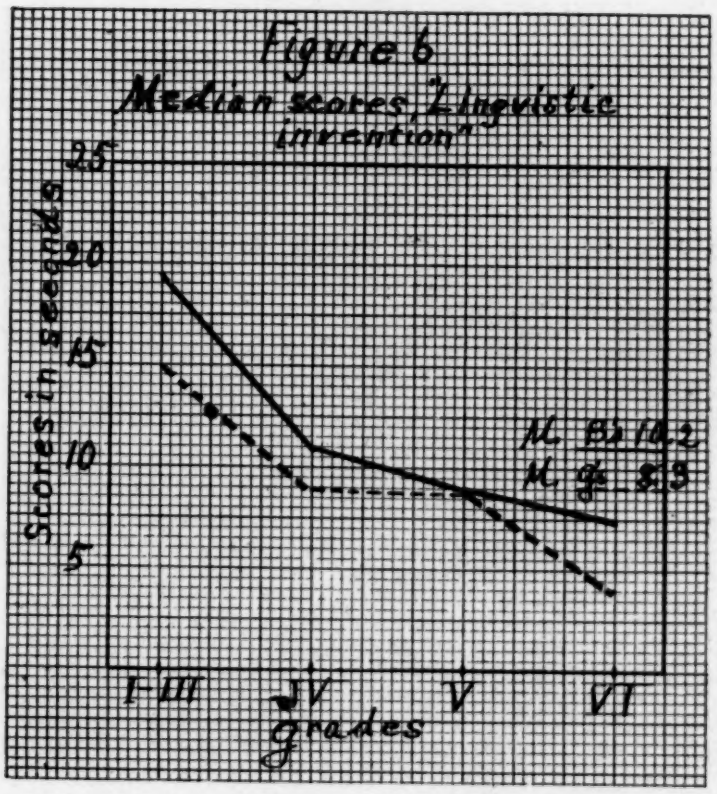
Figure 6
Median scores, Linguistic invention
Scores in seconds
25
20
15
10
5
M. Bs 10.2
M. Gr. 5.3
I-III
IV
V
VI
Grades

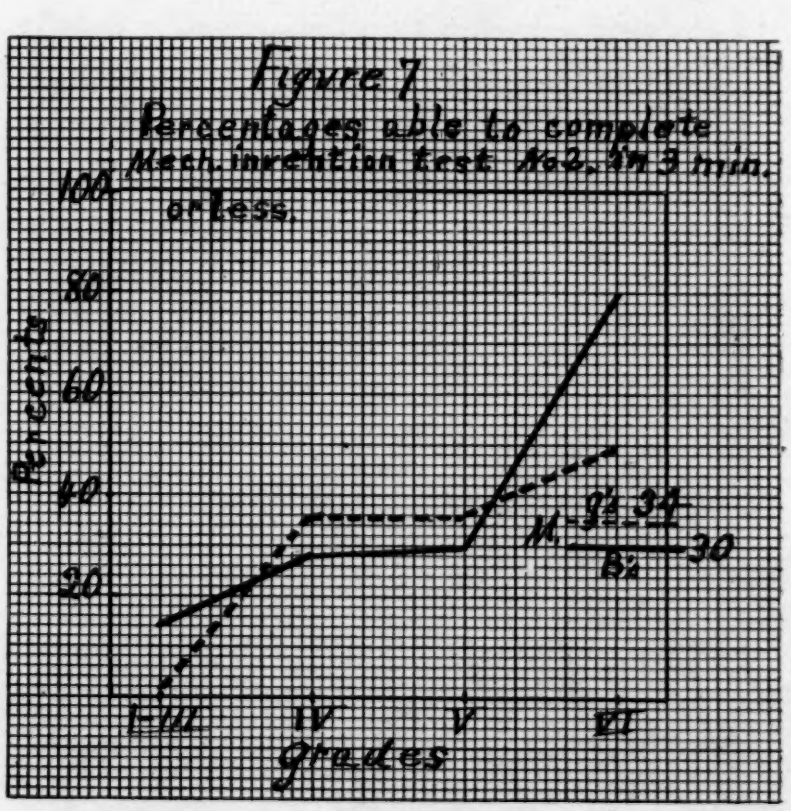
Figure 7
Percentages able to complete Mech. invention test No2. in 3 min. or less.
Percents
100
80
60
40
20
M. 24 34 30
I-III
IV
V
VI
Grades

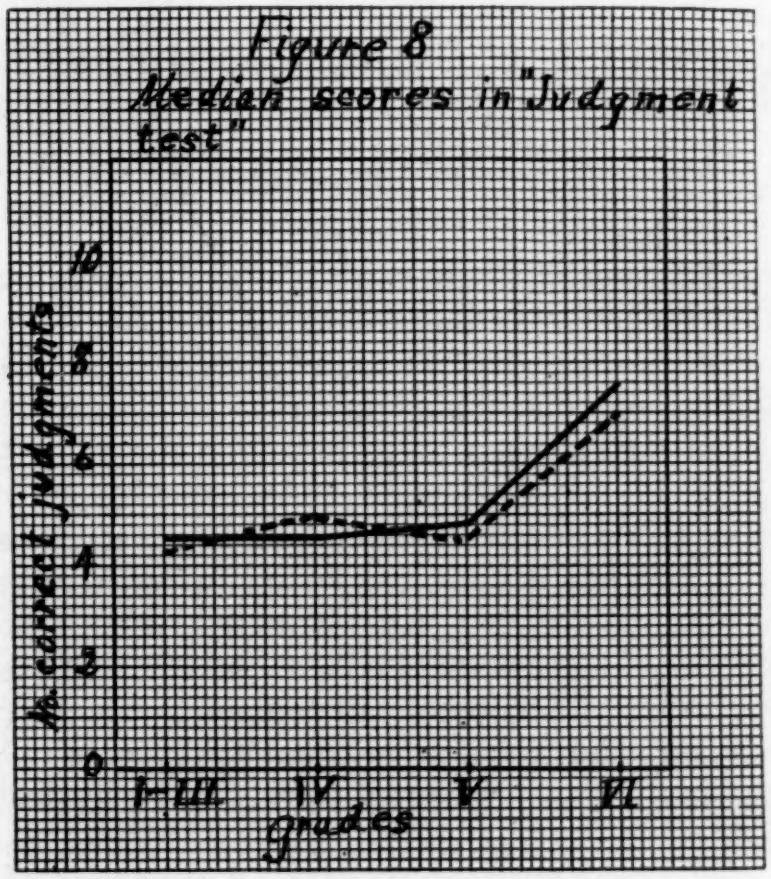
Figure 8
Median scores in "Judgment test"
No. correct answers
10
8
6
4
2
0
I-III
IV
V
VI
grades

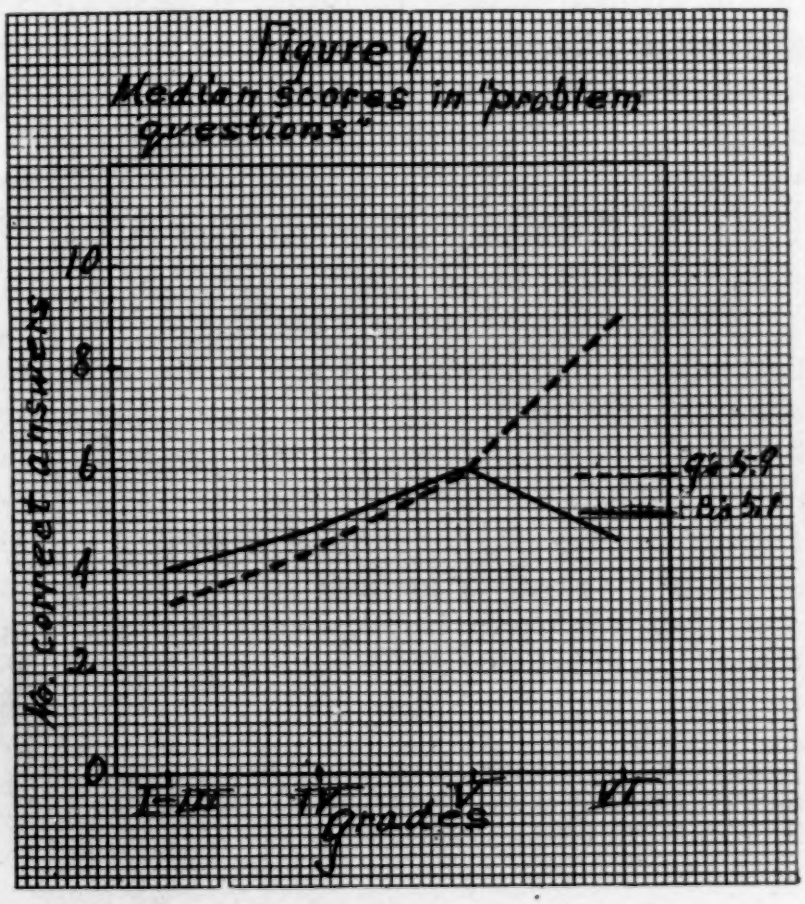
Figure 9
Median scores in "problem questions"
No. correct answers
10
8
6
4
2
0
I-III
IV
Grade V
VI
Gr 5.9
Bx 5.1

A STUDY OF PHYSICAL GROWTH AND SCHOOL STANDING OF BOYS[1]

S. F. STEWART
Rushville, Indiana

This paper represents an attempt to study the relation of physical growth and school standing of boys. It deals with the physical measurements and school standing of 207 boys of the elementary and the high schools of the University of Chicago.

The physical measurements extend over the eight years from 1906–7 to 1913–14. For a few boys the records are complete for as many as seven successive years, but in most cases interruptions occur. The heights are expressed in centimeters and the weights in pounds. These measurements were made by the physical director of the school of education according to generally accepted methods. Each scholarship mark is the average of the office marks for the regular academic subjects. Marks for writing, drawing, music, manual work, and physical exercises are not considered.

The first part of the study deals with average heights and average weights of groups of boys of different school grades. The second part is a study of the individual records of twenty-nine boys whose physical measurements are complete for four or more successive years. The third part is a summary of the points suggested.

Part 1. Average Heights and Weights

The groups considered in the first part include all the boys who have physical measurement records for the eighth grade or for the freshman year in high school. A few of the boys have records for one year only, but in a majority of cases the measurements cover a period of several different years. Chart I indicates that 175 of these boys have records for the freshman year of high school. The groups are smaller for the preceding and the following years. Care is taken to include all of the boys who have records for the freshman year because that is the period of greatest average increase in weight and in height as indicated in Charts III and IV.

[1] This study was made in connection with a course in Experimental Education under Dr. Frank N. Freeman of the University of Chicago. To him the writer is under obligations for suggestions and criticisms.

CHART I

Averages of Weights by Ages and Grades

Ages	Number of Boys	4	5	6	7	8	I	II	III	IV
9	12	3 58.1	7 66.14	2 61.3						
10	26		9 65.3	13 73.27	3 68.3	1 70				
11	38		3 78.2	15 71.4	17 88.4	3 75.9				
12	63		4 79.8	5 91.7	21 79.18	27 95.16	6 90.6			
13	113		1 72	4 89.8	7 91.47	41 90.3	54 96.9	4 88.6	2 95.5	
14	120				6 100.9	15 106.14	73 102.8	25 105.8	1 107.5	
15	97				1 85.2	5 116.5	31 114.8	44 117.9	13 123	3 122.9
16	65					1 94.8	10 126.3	20 132.8	25 125.7	9 135.8
17	36						1 107	8 124.7	2 124.5	25 137.19
18	13									13 142.8
	583	3	24	39	55	93	175	101	43	50

Chart I gives the averages of 583 weights distributed by ages and by school grades. The heavy rectangles indicate the normal positions of those boys who entered school at six years of age and who have advanced one grade each year. The small figure above each average indicates the number of boys whose measurements are represented by that average.

For any given age, the boys one year above the normal grade average heavier than the boys of the normal grade. Also, except for fifteen and for seventeen years, the boys one year below the normal grade average heavier than the boys of the normal grade. Possibly the exceptions at fifteen and seventeen years may be partly due to the changes in relative weight which occur at this time. Some of the boys who were light now surpass some of the boys who were heavier.

Until the eighth grade is reached, the groups of pupils below normal grade are relatively small. During high school, the groups of pupils above normal grade are relatively small. This change in the distribution is due to the fact that more boys are

considered for the freshman grade than for the preceding and the following grades. Probably averages of small groups are not very significant. So emphasis will be placed upon those averages which represent larger numbers of measurements. When this is done, there appears to be a positive relation between weight and scholarship in the elementary grades, but the relation in high school is uncertain.

CHART II

Averages of Heights by Ages and Grades

Ages	Number of Boys	4	5	6	7	8	I	II	III	IV
9	11	3 127.9	6 134.4	2 139.6						
10	26		9 136.2	13 140	3 143.7	1 134.9				
11	38		3 139.9	15 139.27	17 146.8	3 145.7				
12	63		4 146.8	5 146.12	21 145.5	27 152.0	6 152.6			
13	112		1 138	4 152.8	7 152	41 151.3	53 154.8	4 153.8	2 151.5	
14	120				6 156.8	15 159.8	73 159.1	25 160.6	1 162.0	
15	97				1 146.2	5 166.4	31 164.8	44 166.2	13 170.2	3 172.9
16	65					1 156.1	10 168.6	20 170.9	25 170.2	9 174.0
17	36						1 162	8 168.4	2 167.5	25 173.5
18	13									13 172.8
	581	3	23	39	55	93	174	101	43	50

Chart II shows in a similar manner the distribution of the averages of 581 heights. The interpretation is similar to the interpretation of Chart I. For any given age, boys one year above the normal grade average taller than boys of the normal grade. Also, except for 15 and for 17 years, the boys one year below the normal grade average taller than boys of the normal grade.

If medians instead of averages are used in Charts I and II, the distribution will remain unchanged.

Smedley and Porter found that boys at or above grade were heavier and taller than boys below grade. In their investigations each boy was measured only once. Baldwin considers the same

boys through a period of several successive years. His results indicate that height and weight are favorable conditions to scholarship. However, the above investigations do not indicate that boys below normal grade are likely to be taller and heavier than boys of the normal grade.

CHART III

Averages of Increments of Weight by Ages and Grades

Age	Number of Boys	5	6	7	8	I	II	III
10	21	9 5.5	10 9.3	2 6.9				
11	36	2 7.3	14 7.7	18 9.9	2 9.7			
12	47		3 8.2	23 9.9	21 11.0			
13	62			7 14.4	34 10.1	21 14.8		
14	54				8 14.7	41 14.6	5 22.1	
15	29					15 12.2	14 10.7	
16	17					5 8.6	3 9.6	9 10.2
	266	11	27	50	65	82	22	9

CHART IV

Averages of Increments of Height by Ages and Grades

Age	Number of Boys	5	6	7	8	I	II	III
10	20	8 3.3	10 5.6	2 4.3				
11	36	2 4.9	14 5.5	18 5.6	2 5.2			
12	46		3 8.2	23 5.57	20 5.43			
13	61			7 8.14	34 6.51	20 7.8		
14	54				8 6.9	41 7.4	5 9.8	
15	29					15 6.25	14 3.68	
16	18					5 2.6	3 2.8	10 3.55
	264	10	27	50	64	81	22	10

Charts III and IV show the distribution of the averages of 266 increments of weight and 264 increments of height. With one exception, groups one year ahead of the normal grade are increasing more rapidly both in weight and in height than are normal groups of the same age. For some ages, the groups one year below normal are increasing more rapidly than are the groups of normal grade.

So when averages of groups are considered, the preceding charts suggest that until near the middle of high school most boys superior in weight and in height are likely to excel in school standing. But we find numerous exceptions to the above suggestion. Group averages are not trustworthy in that they do not provide for individual variations. For example, of the freshman boys who were thirteen years of age, the heaviest boy weighed 199 pounds while the lightest weighed 73.9 pounds. The group average was 96.9 pounds. Of the eighth grade boys who were thirteen years of age, the heaviest weighed 126 pounds while the lightest weighed 68.3 pounds. The group average was 90.3 pounds. Similar cases of overlapping occur throughout the four charts

CHART V

Averages of Weights by Ages and Grades

Age	Number of Boys	5	6	7	8	I	II	III	IV
9	6	5 69.8	1 61.6						
10	15	5 68	9 75.4	1 68.3					
11	19	2 85.5	7 72.5	9 84.6	1 73.2				
12	21		3 87.7	7 80.1	10 91.6	1 89.2			
13	24			4 91.7	8 87.9	12 103.2			
14	25				4 105.8	13 107.2	8 112.		
15	13					3 111.2	8 121.2	2 137	
16	13						3 123.6	8 130.3	2 149
17	9							1 125	8 140.1
18	1								1 132
	146	12	20	21	23	29	19	11	11

CHART VI

Averages of Heights by Ages and Grades

Age	Number of Boys	5	6	7	8	I	II	III	IV
9	6	5 137.9	1 138.3						
10	15	5 137.7	9 140.8	1 141.2					
11	19	2 142	7 139.8	9 146.1	1 145.1				
12	21		3 144.2	7 144.9	10 152	1 151.9			
13	24			4 141.0	8 145.8	12 144.1			
14	25				4 157.9	13 159.6	8 164.2		
15	13					3 161.9	8 168.9	2 174.7	
16	13						3 174.0	8 171.5	2 182.6
17	9							1 175	8 174.6
18	1								1 177.3
	146	12	20	21	23	29	19	11	11

given above. Hence to apply group averages to individual pupils would be misleading.

That group averages are not sufficient will be shown by a comparison of the results secured when the 29 boys who have records for several successive years are considered in groups and when they are considered as individuals.

Chart V distributes the averages of 146 weights of the 29 boys. Chart VI distributes the averages of 146 heights of the same group. Twelve of these 29 boys have records extending back to the fifth school grade. Three of the boys have records for seven years and three have records for eight years in succession. These 29 boys were members of three different classes. Eleven were members of the class of 1914, fourteen were members of the class of 1916, and four were members of the class of 1917. These charts are to be interpreted in a manner similar to Charts I and II. Chart V indicates that all of the groups above normal, also all of the groups below normal and under fourteen years of age, average heavier than the normal groups. Chart VI indicates

that all but the thirteen year old group above normal, and the eleven, sixteen, and seventeen year old groups below normal, average taller than the normal groups.

CHART VII

Group of Eight Boys Divided on Basis of Scholarship

	Freshman	Sophomore	Junior	Senior
Upper Group Boys	B, C, D, I	B, C, D, I	B, E, G, I	A, B, E, I
Average Wt.	109.4	122.5	125.8	146.4
Average Ht.	161.6	171	172.6	176.0
Lower Group Boys	A, E, F, G	A, E, F, G	A, C, D, F	C, D, F, G
Average Wt.	106.4	119.8	135	133.9
Average Ht.	160.2	166.9	170.4	175.7

Chart VII concerns eight boys of the same age who entered high school in September, 1910, and who graduated in June, 1914. Since in each class there is considerable variation in age, this is the largest group of the same age and of the same class which has complete physical measurements during the high school course. Each capital letter indicates the same boy throughout the study.

For each year above the elementary school, these eight boys are separated into an upper four and a lower four on the basis of their scholarship marks. Then the averages of height and of weight of the two groups are compared. Throughout the four years the half which ranked higher in scholarship also averaged taller. And, except in the Junior year, they also averaged heavier. The exception in the Junior year is due partly to the fact that E and G now enter the upper scholarship group and displace C and D who are slightly heavier. These four boys are so nearly equal in scholarship that a very slight change may rearrange the ranking.

Thus it is seen that a study of group averages of the twenty-nine boys, as indicated in Charts V and VI, shows practically the same thing as does a study of the group averages of the entire 207 boys as indicated in Charts I and II. That is, groups one year above the normal grade average both heavier and taller than groups of the normal grade. Also, in some cases, groups one year below the normal grade average both heavier and taller than groups of the normal grade. Chart VII indicates that in a group of boys of the same age and the same class, those of high scholarship average both heavier and taller than those of low scholarship.

Part II

Next, the twenty-nine boys are considered as individuals. In Chart VIII (which is too voluminous to print in this article) a comparison of scholarship marks with increments of weight and height is made for the different years. Curves are plotted to represent the scholarship marks and the increments of growth for each of the twenty-nine boys. Since these boys were members of three different grades or classes, the average of the teachers' marks for each group is considered separately. For example, Boy A was one of the eleven boys of the class of 1914. So his scholarship mark for any year is considered with reference to the average of the teachers' marks for the eleven boys of that year. All of the eleven boys did not have exactly the same subjects, and of those who did have the same subjects probably all did not have them under the same teachers. But it is probable that the teacher's marks for any group of boys are a fairly reliable indication of the relative ability of the members of that group.

Since standards of grading may vary from year to year, a common basis of comparison is desirable. So the average of the teachers' marks for each year is weighted by being raised to 100% and is then represented by the base line. Then the average scholarship mark of each boy for the year is weighted in the same degree. For example, an individual average that was nine-tenths of the class average now appears as 90%, and is located ten points below the base line. An individual average that was eleven-tenths of the class average is now 110%, and is located ten points above the base line. For each group, the average increments of weight and of height are found. These averages are represented by the base line. Then the difference between the average and any individual increment is indicated by the distance of the curve from the base line. For example, if the average increment of height is seven centimeters, this average is represented by the base line. Then an individual increment of twelve centimeters is five centimeters above the average, and is located five points above the base line. In a similar manner, an increment of two centimeters would be located five points below the base line.

In case a change in an increment is accompanied by a corresponding change in the scholarship grade, it is considered a case of positive correlation. In case a change in increment is accompanied by an opposite change in the scholarship grade, the cor-

relation is regarded as negative. If one of the curves changes while the other remains unchanged, the correlation is considered as zero. If different kinds of correlation exist between a pair of curves, and no one kind is prominent, it is regarded as a mixed relation.

Chart VIII, too elaborate to be printed here, indicates the different relations which exist between scholarship and increment of growth. For example, Boy F has a positive correlation between scholarship and weight increment for three years and between scholarship and height increment for one year only. Boy J has a negative correlation between scholarship and height increment for three years, but scholarship and weight increment are negatively correlated during one year only. For Boy U the scholarship and the height increment have a mixed relation.

For the twenty-nine sets of curves, growth increment and scholarship are related as follows:

	Entirely +	Entirely —	Chiefly +	Chiefly —	Mixed or Zero
Wt. Increment and Scholarship	8	6	4	6	5
Ht. Increment and Scholarship	3	3	6	5	11

Thus for the twenty-nine boys, increase in weight and increase in scholarship are positively correlated throughout for eight boys and are negatively correlated throughout for six boys. For five boys the relations of scholarship and weight increment cannot be classified under any one of the four headings. So they are considered as "Mixed." The relations of height increment and scholarship are to be interpreted in a similar manner. The above tabulation suggests that when sets of curves are considered there does not seem to be any decided relation between scholarship and physical growth.

Chart IX distributes by school grades the yearly correlations of all of the twenty-nine boys. The numbers indicate that there is slightly more positive correlation of scholarship and weight increment during high school than during the preceding grades. For scholarship and height increment there is slightly more negative than positive correlation.

Chart X shows the same yearly correlations distributed by ages of the boys. For boys of thirteen years or over, there is slightly more positive correlation of scholarship and weight increment

CHART IX

Correlation of Curves for Each School Grade

School Grades	Weight and Scholarship			Height and Scholarship		
	+	—	0	+	—	0
5–6	2	1		2		1
6–7	8	7	3	6	6	6
7–8	7	6	5	5	8	4
8–I	6	10	1	7	8	1
I–II	6	4	1	3	6	2
II–III	7	3	1	4	4	3
Totals	36	31	11	27	32	17

CHART X

Correlation of Curves for Each Year of Age

Ages	Weight and Scholarship			Height and Scholarship		
	+	—	0	+	—	0
10		1				1
11	5	5		3	4	3
12	4	5	5	2	4	7
13	6	7	3	6	7	2
14	6	3	3	8	3	1
15	6	6	2	3	7	4
16	5	3	2	5	3	2
17		1				1
Totals	32	31	15	27	28	21

than there is for younger boys. There seems to be little relation between scholarship and height increment.

The preceding charts (IX and X) which deal with correlations do not indicate that there is much relation of scholarship and physical growth. But these correlations do not regard the stage of development of the boy. They assume that growth at one period means the same as growth at any other period; also, the size of the boy, or the sum total of previous growth, is not taken into consideration.

The size of the boy and the rate and the time of growth are indicated on Chart XI (too extensive to be printed here). The curves of absolute height and absolute weight of all of the twenty-nine boys are plotted. By these curves, one is able to tell the increment of growth, and the absolute height and weight of each boy. For many of the boys we are able to judge something of the

stage of development. That is, has the boy nearly reached his adult height, is he now in a rapidly growing period, or will the later growth probably be more rapid? In all but six cases, each boy's stage of rapid growth can be approximately located for either weight or height, or both.

When the curves of a group of boys are plotted on one sheet, it is easy to rank the boys according to weight and height during any school grade or at any particular age. Since all of the twenty-nine boys considered have records for the freshman year, the ranking is made for that year. The twenty-nine boys were members of three different classes. Of one class, all but two boys were fourteen years of age during the freshman year. For the thirteen year old boy, the sophomore measurements were used, and for the fifteen year old boy the eighth grade measurements were used. Of the other two groups, a majority were either thirteen or fourteen years of age while in their freshman year. So the weights and heights of each boy at fourteen years of age were used.

CHART XII

Relation of Weight and Scholarship

14–year Weight	Period of Rapid Growth	Kind of Work Done by Each Boy
Heavy	Early	B—high and steady; O—below to above average; S—high; BB—from below to above average; Q—above average and increasing.
Heavy	Medium	C—from high to low; J—average and unsteady; AA—average to high; U—average to low.
Heavy	Late	A—below average and decreasing; X—low.
Average	Early	L—high and increasing; W—below-above-below.
Average	Medium	D—below average and steady; K—high and increasing; H—above average and unsteady.
Average	Late	None.
Average	Indefinite	P—low and unsteady; T—low and decreasing; Z—high and increasing.
Light	Early	F—low and decreasing.
Light	Medium	E—low and unsteady; G—from below average to average.
Light	Late	I—average to high; Y—low but increasing; CC—average to very low; (Y and CC are older than the other boys).
Light	Indefinite	M—high and decreasing; N—high and unsteady; R—above average and steady; V—below average and unsteady.

Chart XII indicates the weight at fourteen years, the period of most rapid increase in weight, and the scholarship standing of each of the twenty-nine boys. The chart may be summarized as follows:

Heavy boys who develop early, average boys who develop at medium periods, and light boys who develop late, rank well in scholarship in eight cases and poor in three cases. Two of these are boys of over age. Light boys whose rapid development periods can not be definitely located rank well in three cases and low in one case. These indefinite cases seem to have no rapid growth previous to fourteen or fifteen years. It is probable that the rapid period occurs late if at all. Heavy boys who develop late and light boys who develop at early or medium periods rank well in no case and poor in four cases. For the remaining cases, it is difficult to make any general statement. Four are doing poor work, three are doing good work, and the work of the other two is about average.

CHART XIII

Relation of Height and Scholarship

14-year Height	Period of Rapid Growth	Kind of Work Done by Each Boy
Tall	Early	B—high and steady; S—high.
Tall	Medium	A—below average and decreasing; C—from high to low; L—high and increasing; Q—above average and increasing.
Tall	Late	J—average and unsteady.
Tall	Indefinite	O—below to above average; AA—average to high.
Average	Early	U—average to low.
Average	Medium	K—high and increasing; X—low; BB—from below to above average.
Average	Late	G—from below average to average; H—above average and unsteady.
Average	Indefinite	N—high and unsteady; P—low and unsteady T—low and decreasing.
Short	Early	None.
Short	Medium	F—low and decreasing; I—average to high; V—below average and unsteady; Y—low but increasing.
Short	Late	D—below average and steady; E—low and unsteady; CC—average to very low.
Short	Indefinite	M—high but decreasing; R—above average and steady; W—below-above-below.
Indef.	Indefinite	Z—high and increasing.

Chart XIII suggests the following relations between scholarship and height:

Tall boys who increase rapidly at early or medium periods and average boys who increase rapidly at medium periods are doing good work in six cases and poor work in three cases. Tall boys and short boys whose periods of rapid increase are indefinite are doing well in four cases and average in one case. It is possible that some of these indefinite cases increase rapidly too late to be indicated on the curves. Short boys of medium period and short boys of late period do well in one case and poorly in six cases. Of the remaining eight cases, three are doing good work, three are doing poor work, and two are doing average work.

Conclusions

From the preceding study of physical growth and school standing, a few points are suggested. Whether such points will hold when large numbers of boys are considered is a topic for further investigation. But for the boys included in this study the following statements seem to be justified:

1. When we consider averages of groups of the same age, the group one year ahead of the normal grade averages both heavier and taller than the group of the normal grade. In some cases the group one year below the normal averages both heavier and taller than the group of the normal grade.

2. When individual curves and correlations are considered without reference to the size of the boy or to his stage of development, it is difficult to see any relation between physical growth and school standing.

3. When individual curves and correlations are considered, together with the size of the body at fourteen years of age and his stage of development, the following are suggested:

a. Heavy or tall boys of early development rank better than heavy or tall boys of late development.

b. Light boys of late development rank better than light boys of early or medium development. Short boys of late development do not rank high.

c. Boys of medium size or of medium period of development are hard to classify, though a majority of them appear to be doing school work of medium rank.

COMMUNICATIONS AND DISCUSSIONS

FURTHER SUGGESTIONS REGARDING MENTALITY TESTS

I have tested several thousand children during the past two years, employing the Binet test and many other mental and educational tests, and I wish to record some of the convictions that have grown out of my experience. In the first place, I agree quite heartily with Professor Yerkes that the age arrangement or method of scoring is not so desirable for scientific use as a method of scoring which gives cumulative credit for each additional element of the test accomplished. Giving an additional credit for each successive task performed will tend, however, to distort the facts as to the relative achievements of two individuals unless these tasks are so selected as to fall at practically equal intervals along a scientific scale for difficulty of performance. If each X in the line shown below represents a task to be performed, and its position in the line represents its location upon an accurate scale for difficulty, the facts regarding relative performances will not be greatly distorted by giving one credit for each task performed.

X X X X X X

Easy Hard

If the same method of scoring were used in the case of the tasks represented below, however, great misunderstanding would result.

X X XXX X

Easy Hard

To say that Individual A had performed three tasks, while Individual B had performed five, would give an incorrect idea of the relative achievements of A and B in this latter instance, although such a statement would be fairly accurate if the tasks were separated as to difficulty by equal intervals, as shown in the first illustration. We must therefore take care, in arranging tasks in series, to know that the differences in difficulty between consecutive tests are relatively equal.

The problem now resolves itself into one of measuring accurately the differences in difficulty between various elements of a test series. I know of no other satisfactory way to determine the relative difficulty

of two such elements of a test than to try both upon a large number of appropriate individuals. The amount of difference needs, however, to be expressed in such terms as will apply equally well to the difference in difficulty between any other two elements of the test, and in such terms that any other investigator may verify the amount by trying the two elements upon a group of individuals chosen in the same manner as the group originally tested.

In certain educational tests we have assumed that ability to accomplish tasks was distributed within a given school grade according to the "normal surface of frequency," and that one such grade was just as variable as any other, because selected in the same manner by equally competent teachers. Some measure of the variability of a school grade group has been used, therefore, as a measure of differences in the difficulty of elements in the test series. It seems probable that some such function of the normal surface of frequency will be used as the common measure of difficulty for tests applicable to school children and young people. For infants, however, we shall need to find some other measure of differences.

The strongest conviction that has come to me in my work with tests is that we must have tests to measure particular abilities, rather than tests of general mental ability. Instead of reporting a child as mentally defective, we must point out the respects in which he is deficient, and to what degree. A physician does not diagnose a case as merely "a serious illness," but he seeks to discover the respects in which the patient is ill, and to base his treatment upon these particular symptoms rather than upon "poor health in general." Psychologists and educators must likewise be able to point out in what respects a boy will probably be defective for life, and must then seek to adjust training to the individual's possibilities.

As an illustration of the above statements, I found fifty boys in certain institutions of New York State who by the Binet tests were clearly three years or more below normal children of their age. A language completion test applied to these same fifty boys showed that forty-eight of them were three years or more below the normal child of the same age in their ability to complete incomplete sentences. It would have been unfair, however, to have condemned these boys as totally unpromising, for a test of mechanical construction ability applied to them showed that thirteen of the fifty were able to accomplish as much or more than the average child of the same age in the New York City public schools. It is true, of course, that good qualities go

together, and that ability to do mechanical construction work is probably closely related to ability to succeed in the Binet tests; but what we need, it seems to me, is a large variety of tests along many different lines, each test being carefully graded to show how difficult a task of that particular kind the child can accomplish.

M. R. TRABUE.

Teachers College, Columbia University.

Within the past few years one phase of my official duties has involved the applying of mental tests to several hundred subjects ranging from kindergarten children to college students. About 150 of these subjects were tested individually by me, over 100 cases involving my applying the Binet-Simon scale. Within this time also, the training and constant supervision of some 120 Binet examiners has devolved upon me. The following opinions are the result of this experience.

The criticism of the Binet-Simon test methods appears to have been due in part to the following causes: (1) The failure to distinguish clearly between the ability to administer the tests scientifically and the ability to interpret the results of such tests in the light of established or partly established norms, clinical data, etc. Graduate students have been turned loose with a syllabus-manual as practically certified Binet testers. They have registered so-called mental ages which educators, jurists, and others have been left to accept as true mental ages without regard to racial or social norms.

(2) The over-idealism of those who have been hoping (a) to establish a universal scale of mentality tests such as Yerkes heralded last year but has since abandoned, or (b) a national American scale such as Terman has been arranging.

(3) Also another form of over-idealism which prohibits interest in a provisional uniform method because statistical data are lacking as to the correctness of each detail of procedure.

(4) The absence of a laboratory manual of the Binet-Simon tests which attempts to answer provisionally the hundred and one questions of procedure in matters of test order, second chances, uniform scoring, etc., involved in every Binet examination. Yerkes' manual is an advance in this direction for those who are willing to abandon the mental level principle.

(5) The resulting lack of uniformity among mentality examiners. This was shown to be of a serious nature by an experiment conducted in New York some years ago and reported to me by Miss Farrell, Inspector of Ungraded Classes.

The value for preliminary classification of the mental level principle has been accepted either in published statements or in actual practice by (1) such experimentalists as Stern, Terman, Warren, Whipple, Witmer and others. (2) Such authorities on abnormal psychology as Goddard, Rosanoff, White and others. (3) Such educational executives as Superintendent Cornman of Philadelphia, Superintendent Hoke of Richmond, and others. (4) Such leaders in child welfare as Judge Brown of Cleveland and a large number of others who are in sympathy with the movement to use applied psychology in helping to solve social problems.

In the light of the above considerations I would suggest (1) That we endeavor to define more accurately the qualifications in the matter of training and experience for (a) Examiner psychologists who should know both how to administer accurately the several mentality scales that are now in use and how to interpret the total age scores and the individual pluses, minuses or points in the light of the racial and social averages obtained to date, together with all the other educational and social data which should accompany each test result; and (b) Field Examiners or Binet assistants, advanced students in psychology and education who should probably have as a minimum 60 hours of laboratory work together with courses in child psychology and in the technique of applying one of the accepted scales with the aid of a laboratory manual that does not necessitate psychological improvisation.

(2) That we endeavor to educate our present and future mentality examiners to the point of realizing that a Binet age, a Goddard age or a Terman age must be translated into a true mental age except for children of the same race and opportunity as those upon whom the standardization has been based. I trust that Professor Terman will emphasize this point and the need of special experience in interpreting results in connection with his revision.

(3) That we conserve what is best in the Binet principles and methods of gauging pre-adolescent mentality. For example (a) Binet after long experience found that in gauging pre-adolescent mentality, the ability to comprehend and act upon uniform directions without further explanation or demonstration formed a most important element in the complex mental processes being tested throughout the

scale. Yerkes appears to have abandoned this idea in his point-scale method by warning us not to proceed with a test until we are sure that the subject understands the instructions.

(b) Similarly, Binet found that one of the most valuable factors in his scale consisted in presenting the identical pictorial and verbal situations to all subjects so that he might compare their types of response. Terman is using a different set of pictures from those used either by Binet or by Goddard and others. Moreover, in the last form of his arrangement that I have seen Termar appears to prescribe a different form of question according as you guess the subject to be mentally three, seven or fifteen.

(4) That we coöperate in perfecting the Binet principles and methods by continuing the valuable work begun by the Informal Conference on Binet Testing at Buffalo. With this view in mind I am about to present to Professor Warren, under whom I began this psychodetic work as his Research Assistant, a provisional uniform method of applying the Binet-Simon tests as regards test order, details of administering and scoring each test, etc. This method endeavors to follow as closely as possible both the spirit and the letter of the Buffalo Conference. The method is being used under my supervision in the Philadelphia Public Schools. A Laboratory Manual embodying this method is about to be printed by the Trade School for use under my personal direction.

There are a number of other considerations which I have been hoping to embody in an article on the subject when time permitted, as for example the effect of wide range testing coupled with scientific interpretation in minimizing the importance of an ideal placing of the tests on the basis of universal or of national data, the need of a rational conservatism in applied psychology because of the practical and social issues involved, and similar topics.

NORBERT J. MELVILLE.
School of Pedagogy of the City of Philadelphia.

It has seemed to me that what we need as a basis for constructing scales of measurement is a kind of survey of the various capacities of mind and body which are now open to experimentation by fairly well recognized methods, with a view to determining what the limits of variation of each capacity are at each age. We have in this labatory undertaken the task of making such a survey for the ages from fourteen to eighteen years. When the series is complete we shall have

measures for from 1000 to 1500 individuals of each of those ages in a variety of tests covering motor ability (strength, speed, steadiness, coördination of movement) and mental ability (perceptual processes, memory of several types, learning ability, association, linguistic ability, and ingenuity). The subjects have been chosen in such a way as to represent all native born, white individuals of a given age in the community, except those recognized as feeble-minded. The results for each test are being arranged in a ten-percentile scale. Each person in the series is then having his records in each test marked from one to ten according to the percentile within which his results fall. These percentile ranks are then being averaged first for tests of motor capacity, and second for tests of mental capacity. These average ranks are then to be arranged in a final scale.

In the laboratory in New York, established by Mrs. Lucy Sprague Mitchell, and directed during its first year by Prof. Joseph Hayes of the University of Chicago, similar norms for many of the same tests and some additional ones are being established for the ages from eight to thirteen years. When this plan is successfully carried out for the ages from six to eighteen, we can tell how much overlapping there is from year to year in each type of performance, and have a more adequate basis for deciding just what would be the best method of establishing a satisfactory scale. We could obtain rankings of this sort, for instance—This boy of fourteen ranks in his mental tests as well as twenty-five per cent. of boys of his age. Twenty-five per cent. rank for age fourteen corresponds to median for age eleven. We could do the same with tests of motor capacity. An analysis of each average percentile rank could state whether any separate tests showed marked deviation from its median.

By studying the degree of correlation between tests of different types, we ought to be able to come to some conclusion as to the extent to which abilities are general, and the extent to which they are specific. The number and variety of tests whose records should be included in forming a general scale, the choice of a sets of tests to be used for specific purposes and the weighting of given tests for special purposes are problems that seem open to solution on such a basis.

This plan presupposes the choice of a method of giving and of evaluating each test which is sufficiently satisfactory, and sufficiently well recognized to be used by everyone for purposes of general measurement. In undertaking the work I had thought that this was a safe assumption for certain tests, though I realize that there may be

a serious difference of opinion on this point. At any rate, I would like to take this opportunity to emphasize the fact that numerical mental or physical norms to be of any service must be accompanied by very complete and explicit descriptions of methods of giving and evaluating the tests on which they are based.

This office expects to furnish a scale of the type described for each age from fourteen to eighteen, to furnish a comparison of ranking in this scale with ranking on the Yerkes-Bridges Point scale, and to formulate the relation between rank in mental and motor tests, and subsequent industrial career.

HELEN T. WOOLLEY.
Bureau of Vocational Guidance, Cincinnati Public Schools.

ABSTRACTS AND REVIEWS

WILLIAM HEALY, A.B., M.D. *The Individual Delinquent: A Textbook of Diagnosis and Prognosis for All Concerned in Understanding Offenders.* Boston: Little, Brown and Co., 1915. Pp. xvi + 830. $5.00.

Not only in respect to the timeliness of its subject-matter, but also in respect to the scope and thoroughness of its treatment, its many evidences of painstaking research and its indubitable scholarship, is this book a contribution of the first order.

As our readers will recall, the Juvenile Psychopathic Institute, projected by Miss Lathrop (now chief of the Children's Bureau at Washington) and endowed for five years by Mrs. Dummer, was established at Chicago in 1909. Dr. Healy served as Director of this Institute and the present volume presents to the public the scientific results of this half-decade of activity.

It would be absurd to attempt within the bounds of a book-review anything like a comprehensive analysis or summary of this work, but its general plan may be indicated. Book I (pp. 1–182), entitled "General Data," comprises an introduction and nine chapters dealing with Orientations, the Individual, the Mental Bases of Delinquency, Working Methods (three chapters), Statistics, and General Conclusions (as to methodology and as to treatment). Book II (pp. 183–788), entitled "Cases, Types, Causative Factors," comprises an introduction and 27 chapters dealing with heredity, developmental conditions, physical abnormalities, stimulants and narcotics, environmental factors, professional criminalism, mental abnormality (epilepsy, sexualism, criminalistic imagery, conflicts and repressions, feeble-mindedness, adolescent psychoses, insanity, pathological lying, stealing and accusation, love of excitement, etc.). The material in Book II is enriched and aptly illustrated by a wealth of detailed cases drawn from the records of 1000 repeated offenders selected from the larger number of cases that have been studied by the Institute.

It is evident that the task set by this investigation, that of unearthing the causes of criminalism, is primarily a problem of characterology. It implies an analysis of the motives of human conduct, and this becomes in considerable measure a problem of individual and differentiated psychology.

The following may be cited as random samples of some of the conclusions reached by Dr. Healy: (1) Intimate study of the individual offender is of fundamental importance. No amount of shrewdness and of familiarity with courtroom cases can replace this direct search for causative factors in each individual. (2) "There is no such thing as an anthropological criminal type." (3) A cut-and-dried classification of offenders "appears at once out of the question," though certain causal types may be differentiated, and, in general, a consideration of the mental life is the most direct way of arriving at the causal factors. (4) In each case the several causative factors, *e. g.*, defective heredity, alcoholism, poor parental control, bad companions, interact in the most intricate manner. (5) Punishment is necessary, but should not harm the offender. (6) It is of the first importance to begin treatment of the offender as early as possible in his career. (7) For effective work, courts need to be reorganized by extending the juvenile court method up to the age of 20 or 21 years and by giving to the court handling the offender direct jurisdiction over the contributing agencies to his offense. (8) "Treatment of the physical ailments and incapacities of the offender is often an absolutely indispensable condition for his moral success." (9) "Educational treatment is essential for many delinquents, even though they be adults," especially by arousing healthy interests and healthy mental content. "The schoolroom, for the delinquent, should be the avenue to higher vocational possibilities, to better recreational resources, to appreciation of right methods of thinking." (10) The environment of the offender is "always to be taken account of as a formative influence, at least until the individual is thoroughly crystallized in his habits." (11) Institutions, *e. g.*, the reformatory, have a real opportunity to supply a corrective environment, now often lost through self-complacency and through lack of adequate classification of offenders and insufficient individualization of treatment. (12) "A great feature of treatment is the careful carrying over of offenders through the period of adolescent instability." (13) "A very weak point in practically all social and moral therapy is the lack of follow-up work." (14) Idiots are rarely, if ever, criminals. Imbeciles are "not at all numerous" in court cases. Morons are more important, especially because they are not, as popularly supposed, merely passive delinquents. It is impossible to lay down any general recommendation for the treatment of the moron group (save, perhaps, that they particularly need to be protected from alcohol): the individual varia-

tions in their capacities are too great. (15) "A certain number of individuals have a special, definite, innate defect in the process of self-control." (16) "If the 'moral imbecile' exists who is free from all other forms of intellectual defect he must indeed be a *rara avis*." (17) The Binet scale "is really a valuable measure of the lower general levels of intelligence" and is, within definite limits, "of the utmost practical usefulness."

By these random samples, together with the brief synopsis of the contents of the volume, I have sought to give a rough notion of the aim and treatment of Dr. Healy's work. I realize the inadequacy of this method of reviewing so compendious a study—which is, I repeat, a contribution of the first order—and I must atone for it by urging psychologists, physicians, school and court authorities and all who aim to understand better the complex and vital problems with which it deals to get acquainted with it at first-hand. G. M. W.

HENRY HERBERT GODDARD. *Feeble-mindedness: Its Causes and Consequences*. New York. The Macmillan Company, 1914. Pp. XII, 599. $4.00.

No one has had better opportunities for extensive and intensive studies of feeble-mindedness than the author of this book. As director of the research laboratory of the Vineland, New Jersey, Training School for the Feeble-minded, he has not only made psychological studies and tests of individual cases, but has been able to repeat these studies at intervals on the same individuals, and, with the assistance of the staff of field workers, has collected a mass of material bearing on the family life of each individual that is unsurpassed in the history of psycho-pathology.

In his introductory chapter the author mentions four types of problems which call for investigation in the study of feeble-mindedness. First, there is the social problem, dealing with the prevalence of feeble-mindedness at the present time, the place the feeble-minded occupy in contemporary life, and their relation to crime, pauperism, intemperance, the social evil, incompetency, and disease. The second is the psychological problem. What sort of mental processes have the feeble-minded? How do they compare with normal individuals in the various aspects of mentality, as perception, discrimination, association, memory, judgment, and the like? Third, the pedagogical problem. To what extent can they be trained? What are the

results of the ordinary means of training? Of what moral training are they susceptible? Fourth, the biological problem. What is the cause of mental defectiveness? What is its physical basis, and what means of prevention can be suggested? The discussions in the present volume have to do primarily with the social and biological problems. The psychological and pedagogical problems are still under investigation.

To obtain reliable data regarding the family history of these unfortunates is a matter requiring no little sagacity, tact, perseverance, and discrimination. Owing to the importance of the subject, a whole chapter is devoted to an account of the manner in which the field material was collected. Chapter three constitutes the body of the book, and comprises almost four hundred pages. There are detailed reports of three hundred and twenty-seven cases. The general plan of the description of these cases is a statement of the patient's present condition, a description of his status in educational work; samples of handwriting and composition in the more advanced cases; a chart of family connections, in so far as these have been ascertained; and such general descriptive data as will contribute to a better understanding of the individual case. The cases are grouped as hereditary, accidental, neuropathic ancestry, and no assignable cause. By far the larger group is the hereditary, comprising in all one hundred and ninety-eight of the three hundred and twenty-seven cases.

This classification is not altogether satisfactory, for a study of the heredity charts and the case histories shows that many of those in the hereditary group have been victims of disease or accident, and many of the accidents enumerated are very doubtful causes of feeble-mindedness. Moreover, those cases ascribed to neuropathic ancestry have developed their present state of defect only under hereditary influences. Chapters four and five, however, which discuss the probable causes of the abnormalities described and present the data in tabular form, show that the principle of classification is a strictly pragmatic one. Those cases in whose families mental defect is clearly and strikingly manifest are classed as hereditary, those whose relatives show no feeble-mindedness, but give evidence of insanity, paralysis, epilepsy, or other pathological nervous conditions, are put in the group with neuropathic ancestry, while those in whose families no nervous disturbances are found, but whose histories show serious accident or disease (particularly meningitis) are attributed to accident. It may be, as the author carefully points out, that the neuropathic

conditions and the feeble-mindedness are both manifestations of a more subtle biological defect, while it is very doubtful whether many of the accidents assigned as the causes of feeble-mindedness are of themselves capable of producing such effects.

Throughout the discussion one can see that the author is strongly impressed by the hereditary data revealed by the field studies, and the last five chapters of the book are devoted to the systematic development of a theory of feeble-mindedness as an inherited unit character in accordance with Mendelian principles. It is true that the individual family is too small to show clearly the three-to-one relation discovered by Mendel in the crossing of hybrids, but if we add together all the offspring from similar matings (as Mendel himself did in the case of peas) and compare the actual findings with the theoretical expectation, we are at once impressed with the closeness of the agreement. But the phenomena characteristic of the development of the feeble-minded are negative rather than positive, *i. e.*, the individual develops normally, or with an approximation to the normal order, up to a certain point, then the process begins to slow down and finally comes to a complete stop, while the normal child goes on to maturity. Accordingly it would seem that normal intelligence is the positive or dominant characteristic, and feeble-mindedness is the negative or recessive aspect. In other words normal intelligence is due to a specific determiner or set of determiners in the germ plasm, while in the feeble-minded these determiners are lacking. When pure feeble-minded strains are crossed, the offspring is always feeble-minded (and for this the author believes he has good evidence), while if normal and feeble-minded strains are crossed the first generation gives no evidence of the defect, but subsequent generations approximate the Mendelian proportions.

What practical effects will such a theory have on our attitude toward the problem of feeble-mindedness? The author is not at all sanguine about the feasibility of stamping it out by legislative enactment. Sterilization, which would undoubtedly be effective if carried out consistently for several generations, meets with strong popular opposition that is so deeply rooted in the prejudices that cluster about our ideas of virility as to offer little hope of improvement. Colonization is a means of procedure that has a greater popular appeal, but we are just beginning to realize how many hundreds of thousands of feeble-minded we have in this country, and the task of segregating these in colonies is appalling. Education has been the hope of well-

meaning people for a generation, but our studies show us that the feeble-minded are educable only in certain limited directions. Rather the author finds the hope of civilization in educating people of normal intelligence to understand the nature and possibilities of the high-grade feeble-minded, to secure his protection in the kind of work he can do, to guard him against the exploitation of the evil-minded, and to utilize him as far as possible in the work of the world.

The book is a masterful presentation of authoritative data that touch upon some of our most vital social problems. Whether the distinction between normal intelligence and feeble-mindedness is sufficiently sharp and definite to warrant the author's conception of it as a unit character may well be disputed. In any case the view is stimulating and suggestive, and should give rise to further careful studies.

J. C. B.

THE
JOURNAL OF EDUCATIONAL PSYCHOLOGY

INCLUDING EXPERIMENTAL PEDAGOGY, CHILD PHYSIOLOGY AND HYGIENE, AND EDUCATIONAL STATISTICS

Edited by

W. C. BAGLEY, C. E. SEASHORE,
J. CARLETON BELL, Managing Editor. GUY MONTROSE WHIPPLE.

Published monthly except July and August by
WARWICK & YORK, INC.,
LANCASTER, PA. BALTIMORE, MD.

Manuscripts for publication, books or other material for review, and news items of interest to educational psychologists should be addressed to The Managing Editor, J. Carleton Bell, 1032A Sterling Place, Brooklyn, N. Y. Research contributions from England may be sent to W. H. Winch, Esq., Melville Villa, Overhill Road, Dulwich, London.

The Journal is issuing a series of Monograph Supplements for the publication of more extensive researches in educational psychology. For details prospective contributors may apply to Guy Montrose Whipple, University of Illinois, Urbana, Ill.

The price of the Journal is $2.50 per year of ten monthly issues. Foreign postage is 20 cents extra for Canada and 30 cents for other countries in the Postal Union.

Subscriptions and correspondence concerning business matters should be sent to Warwick & York, Inc.

VOLUME VII SEPTEMBER, 1916 NUMBER 7

EDITORIAL

In the current discussions of the reorganization of the curriculum it is encouraging to note that increasing emphasis is laid upon the arts.

MUSIC IN THE CURRICULUM

While the trend in this direction is primarily toward the practical and industrial arts, indications are not lacking that the fine arts are destined to receive more attention in the future than they have in the past. If education is to conserve and advance those forms of human activity that are ennobling and elevating and cultural as well as those that are useful and materially productive, then music, painting, and sculpture must take their place with literature as integral elements of the curriculum. For it is in these fields of effort that human nature shows its widest departure from the level of the brutes.

Most schools already devote some time to singing, and there is a pretty general provision for drawing and perhaps a bit of modelling in some of the grades. As yet, however, these subjects are regarded even by school people as extras, requiring special supervisors and in some cases special teachers, and one rarely finds a course in music, for example, in which the subject is worked out in a thorough-going and systematic manner as an essential part of the educational pro-

gram. Yet it would not be difficult to show that there is a higher educational value in a knowledge of the best musical compositions for the voice, the piano and the violin, and in the development of a critical taste for and an appreciation of good music than in nine-tenths of the arithmetic and grammar taught in the schools. Such an understanding of music is not to be developed by mere passive listening to the rendition of selections by others. If our young people are to learn to know music, they must be taught to make music. Any one who can talk and who has a reasonable tonal discrimination can be taught to sing. This capacity could be utilized by arranging the great vocal masterpieces for school use and having them learned by the children not merely as pretty pieces to be sung, but as great works of art which are worth incorporating into their own thinking.

But it is to instrumental music that we must look for much of the educative value of music study,—and this field has been almost wholly neglected by the schools, It is, therefore, a satisfaction to note that the Board of Education of New York City has recently provided that for a full year's participation in either a school choral society or an orchestra two credits shall be allowed in the second, third and fourth years of the high school course. This is still far from offering a well-organized course in instrumental music, but it is a step in the right direction. Another gratifying indication of the same tendency is the bequest of $700,000 to the Chicago Orchestral Association the income of which is to be used in support of a public school of music in connection with the symphony orchestra. The testator expressed the hope that other gifts and bequests would be added by those who recognize "the value of a great symphony orchestra in educating the people to appreciate whatever is great and beautiful." If the public schools would now offer the pupils in the sixth, seventh and eighth grades the opportunity to learn to play some musical instruments, and would provide for more chamber music in the high school, we should make rapid strides in the realization of the educational value of music.

J. C. B.

NOTES AND NEWS

Under the title of "A School of Superintendence" a conference of school superintendents was held last June at Lincoln, Nebraska, under the auspices of the University of Nebraska and the State Department of Public Instruction. One of the principal speakers was Dr. Walter S. Monroe, of Emporia, Kansas, who gave a series of addresses on the following topics: "Measuring the Work of Public Schools," "The Application of Standards in School Supervision," "Educational Diagnosis and Remedial Instruction," and "Educational Research and its Significance to Superintendents."

Mr. W. W. Charters, of the University of Missouri, delivered a series of five lectures on Methods of Teaching before the Graduate School of Agriculture at Amherst, Massachusetts, July 3–8. Mr. W. C. Bagley, of the University of Illinois, lectured upon the Foundations of Pedagogy at the Graduate School of Agriculture during the same week.

Dr. David Spence Hill, who has distinguished himself as director of the Bureau of Educational Research, in New Orleans, was given the honorary degree of Doctor of Laws by the University of Kentucky this summer.

With the moving of the Johns Hopkins University to Homewood the psychology department will be divided. Part will be established under the direction of Professor Knight Dunlap in the academic building, while the other part, under the direction of Professor J. B. Watson and Dr. Lashley, will be moved to the psychological laboratory of the Phipps Psychiatric Clinic at the medical school. Professor Dunlap has been advanced to the rank of full professor.—*School and Society.*

During the spring the Los Angeles board of education appointed a commission consisting of Dr. Albert Shiels, director of the bureau of reference and research, New York City Schools, Professor J. F. Bobbitt, University of Chicago, and Dean Walter A. Jessup, University of Iowa, to make a study of the Los Angeles schools, and report on their future conduct. Since that time Dr. Shiels has been elected superintendent of the Los Angeles schools. Quite recently Dean Jessup was elected president of the University of Iowa.

Dr. William Frederick Book, formerly professor of educational psychology at Indiana University, who for the past three years has been on leave of absence from the University, and employed by the Indiana State Board of Education to organize vocational education in the State under the new Indiana Vocational Education Law, will return to the University next fall, as soon as the state vocational surveys now being made under Dr. Book's direction are finished and recommendations made for further developing vocational instruction in the state. At a recent meeting of the Board of Trustees, Dr. Book was appointed director of the psychological laboratory, in the Department of Sociology, the position formerly held by Professor M. E. Haggerty, and professor of educational psychology in the School of Education.

Charles H. Winslow, of New York City, will have charge of a vocational survey of Indiana conducted under the direction of the State Board of Education, the Indianapolis public schools, and the National Society for the Promotion of Industrial Education. — *School and Society.*

Mr. J. C. Brown, who had charge of the work in administration and supervision in the School of Education, University of Illinois, resigned at the close of the year, 1915–16, to accept the presidency of the St. Cloud, Minnesota, State Normal School.

Mr. W. S. Miller, who was assistant in education and secretary of the School of Education, University of Illinois, from 1912 to 1916 has been appointed to an assistant professorship in education at the University of Minnesota. He will have charge of the practice-school connected with the College of Education.

Mr. B. F. Pittenger has been appointed associate in education at the University of Illinois for the year 1916–17. During the second semester of last year he held a temporary appointment as lecturer in education at the University of Illinois.

A psychological laboratory has recently been established at Bellevue Hospital, in New York City, under the direction of Dr. Menas S. Gregory, chief of the Psychopathic and Alcoholic Services. Facilities will be provided for both clinical and research work. As these services admit about fifteen thousand patients annually, the opportunities for research will be exceptional. Dr. Leta S. Hollingworth,

formerly psychologist in the Department of Public Charities in New York City, has been placed in charge of the laboratory.

Dr. E. E. Rall, professor of education at the University of Tennessee, Knoxville, Tenn., has been elected president of Northwestern College, Napierville, Ill.

At the University of Virginia Associate Professor Hall-Quest has been made full professor of education. Professor Hall-Quest is the author of a forthcoming book on "Supervised Study" to be published by Macmillan.—*School and Society.*

Dr. Walter S. Hunter, of the University of Texas, has been appointed professor of psychology and director of the psychological laboratory at the University of Kansas.

Four graduate students of psychology have been appointed as fellows for the coming year in the Bureau of Salesmanship Research affiliated with the Carnegie Institute of Technology, as follows: Dwight L. Hoopingarner, of the University of Texas; C. P. Stone, University of Minnesota; Russell L. Gould, Columbia University; Edward S. Robinson, University of Cincinnati. In addition to these appointments, Dr. Kurt Th. Friedlaender of San Francisco has received appointment as Honorary Fellow.